RHODES, MATTHEW TAYLOR, Ph.D. Three Essays in Applied Microeconomics using Panel Data. (2013)
Directed by Dr. Stephen Layson. 164 pp.

The first essay of this dissertation examines the substitution effects between Major League Baseball home games and how these substitution effects impact the attendance returns of a doubleheader—a day in which two games are played for the price of one. Specifically, a model of daily baseball attendance based on utility maximizing behavior is developed and then tested using Major League Baseball data from Retrosheet.org for 1938 to 2009. Findings suggest that home games are substitutes for each other and the substitution effects are strongest when home games are played closer in time. Additionally, the substitution effects between single games and nearby doubleheaders are particularly strong. In fact, these substitution effects are strong enough to overwhelm the positive day-of attendance returns of a doubleheader to where the total effect of doubleheaders on season attendance is negative. Lastly, two implications of this latter finding are discussed: 1. Was the widespread use of doubleheaders, particularly from 1938 to 1985, in the team owners' self-interest; and, 2. After properly accounting for substitution effects, how effective are traditional modern-day promotions in increasing season attendance?

The second essay estimates the environmental impact of sporting events by analyzing a collection of small typically geographically isolated cities which host at least one NCAA football team that competes in the Division I Football Bowl Subdivision (FBS) in 2010. Fixed-effects regressions controlling for differences across cities and across months suggest that cities do experience an increase in pollution levels on and around game days relative to non-game days. These marginal increases were largest in November even after controlling for weather and various trends. However, predicted levels were below EPA daily thresholds. Additionally, hypothetical levels required to increase mortality rates from 0% to 1% were three to eight times larger than observed maximum game day pollution levels. Thus, the estimated marginal increases in daily pollution levels experienced by cities as a result of hosting a college football game are not hazardous and are not expected to increase mortality risks.

The third and final essay estimates the demand for beer in the U.S. from 2001 until 2006 using a multinomial logit discrete-choice model of product differentiation. Using grocery store scanner data from Information Resources Incorporated (IRI), demand estimates are used to evaluate the extent to which two new beer brands—Michelob Ultra and Bud Select—attracted new drinkers. This unexplored aspect to new brands has various public health implications regarding the over-consumption of alcohol. For a single market, counterfactual results based on a simulation involving 50,000 hypothetical consumers drawn from a type 1 extreme value distribution suggest sales were overwhelmingly generated by new drinkers as they accounted for 68% of sales of Bud Select and 74% of sales of Michelob Ultra. Additionally, new drinkers of Bud Select preferred larger package sizes—specifically 12-packs over 6-packs—whereas the reverse held for Michelob Ultra. Lastly, a number of suggestions for future work are provided.

THREE ESSAYS IN APPLIED MICROECONOMICS USING PANEL DATA

by

Matthew Taylor Rhodes

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2013

Approved by

____
Committee Chair

To my parents, Dusty and Carol Rhodes.

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________________
Stephen Layson

Committee Members ______________________________
Peter Bearse

____
Jeremy Bray

____
Christopher Swann

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

ACKNOWLEDGMENTS

I would like to acknowledge the effort, support and friendship of my dissertation advisor Dr. Stephen Layson who over the years has offered me valuable advice, encouragement and opportunity. In addition, I would like to thank my committee members for their exemplary service: Dr. Peter Bearse for his assistance, enthusiasm and programming expertise which in turn made my second essay possible; Dr. Jeremy Bray for data access, encouragement and helpful suggestions which in turn made my third essay possible; and, Dr. Christopher Swann for his support, suggestions and valuable advice over many years.

I would like to acknowledge the contributions of Dr. James Roberts and Dr. Andrew Sweeting at Duke University and Dr. Brian McManus at UNC Chapel Hill and thank them for their valuable advice, helpful suggestions and coursework opportunities. At UNC Greensboro, I would like to acknowledge the various contributions by the remaining faculty, staff and all graduate students with particular thanks to Matt Rabbitt who endured sharing an office with me for the past three years and yet was still willing to provide me with helpful suggestions and programming expertise which proved vital to my third essay. Also, I would like to acknowledge the service of Dr. Stuart Allen and thank him for the countless opportunities he has given me.

Lastly, for their support and encouragement, I would like to acknowledge my family and my friends—many of which I met in the great town of Boone, NC. While they generally disliked economics, they never let me give it up. Even in spite of my reassurance that once finished I would still be unable to tell them which stocks to pick and would still be unable to get them a job at the Fed.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

ESSAY 1: SUBSTITUTION EFFECTS BETWEEN MAJOR LEAGUE BASEBALL HOME GAMES AND THE IMPACTS OF DOUBLEHEADERS

1.1 Abstract

We develop a model of daily baseball attendance based on utility maximizing behavior and then test it using Major League Baseball data from 1938-2009.[1] We find that home games are substitutes for each other and the substitution effects are strongest when home games are played closer in time. We also find that the substitution effects between single games and nearby doubleheaders are particularly strong and that the total effect of doubleheaders on season attendance is negative. This leads us to question whether the widespread use of doubleheaders from 1938 to 1985 was in the team owners' self-interest.

1.2 Introduction

This paper uses daily Major League Baseball (MLB) data from 1938-2009 from Retrosheet.org[2] to estimate how daily attendance at MLB games is affected by the availability of other home games within a short time span. Because MLB games played on nearby days in the same ballpark are likely to be close substitutes, we expect that there is diminishing marginal utility from attending nearby MLB games.

[1] Co-authored with Stephen K. Layson.

[2] The information used here was obtained free of charge from and is copyrighted by Retrosheet. Interested parties may contact Retrosheet at www.retrosheet.org. Lastly, the data were downloaded on July 16th, 2010.

This may explain why current home-stands in MLB are typically only six or seven games in length.[3] Although we do not attempt in this paper to determine the optimal length of a home-stand, our findings are certainly relevant to this question.

In our empirical analysis of daily attendance, we distinguish between attendance at single games and attendance at doubleheaders. Our definition of a doubleheader is two consecutive baseball games played on one day for the same ticket price as a single game.[4] Although doubleheaders are rare in present times, they were common in MLB from 1926 to 1967. In each year of this 41 year period, doubleheaders accounted for at least 20% of the total games played in MLB. The peak year for doubleheaders was 1945 when they accounted for 48% of MLB games. Because doubleheaders are two games for the price of one, attendance at doubleheaders should be higher than attendance at single games, ceteris paribus. Doubleheaders, however, should also have negative effects on attendance at nearby games. Because of diminishing marginal utility, fans should be less likely to attend games (singles or doubleheaders) shortly before or after a doubleheader. Also, doubleheaders should have negative effects on nearby single game attendance as fans rationally substitute doubleheaders for single games.

We develop a model of daily baseball attendance based on utility maximizing behavior that incorporates these types of substitution effects between nearby single games and doubleheaders and then test it using daily MLB attendance data. We divide our entire sample into three sub-periods: 1938-1947, 1953-1984 and 1987-2009. In the first two sub-periods, 1938-1947 and 1953-1984, we find, after controlling for the

[3]This is based on our analysis of modern day schedules via espn.com and annual volumes of the AL Red Books and NL Green Books from 1986-2009.

[4]By doubleheaders we mean single-priced doubleheaders as opposed to separate-priced doubleheaders which are two baseball games played on the same day with separate ticket admissions. We treat separate priced doubleheaders as two single games played on the same day.

impacts of neighboring games, that doubleheaders increase attendance on the days of doubleheaders by 3,713 and 4,177, respectively. We also find that the attendance of a given doubleheader is lowered when a neighboring doubleheader occurs one day after or one day before. Similarly, the attendance at a given single game is lowered when a neighboring doubleheader occurs up to three days after or three days before. With regard to the substitution effects between nearby single games, we find in the first two sub-periods that single game attendance on a given day is negatively affected by the availability of nearby single games one day after or one day before.

In the final sub-period from 1987-2009 when there was a sharp decline in the use of doubleheaders, we no longer find that doubleheaders increased attendance on the days of doubleheaders. Nevertheless, we continue to find that the attendance of a given single game is substantially lowered when neighbored by doubleheaders. We regard this as evidence for our hypothesis of diminishing marginal utility between nearby games. In terms of the substitution effects between nearby single games, we find strong evidence that the attendance level of a given single game is lowered by neighboring single games which occur within seven days.

In terms of the previous literature, only four studies to date–Siegfried and Eisenberg (1980a, 1980b), Hill, Madura and Zuber (1982), Marcum and Greenstein (1985) and Bruggink and Eaton (1996)–examine the relationship between attendance and doubleheaders in the broader context of examining the determinants of baseball attendance. Siegfried and Eisenberg (1980a, 1980b), using annual minor league data from 1973 to 1977, find that the annual number of team doubleheaders had no statistically significant effect on season team attendance. The other three studies use daily MLB data for selected single years and only estimate how doubleheaders affect attendance on the days of doubleheaders.[5] However, none of these studies examined the substitution effect between doubleheaders and nearby games.

Additionally, since doubleheaders offer fans two games for the price of one, doubleheaders can be interpreted as a promotion designed to boost attendance. Thus, our analysis of examining the substitution effects between doubleheaders and nearby games can be easily adjusted to examine the possible substitution effects between promoted games and nearby games. Such an examination would extend the recent literature on promotions since the current empirical work has only examined the attendance impact of promotions on the days of promotions.[6]

We use our daily results and our strong findings of substitution effects between doubleheaders and neighboring games to argue that the total effect of a doubleheader—the impact of doubleheaders on season attendance—for all sub-periods is negative. As a robustness check, we follow the methodology of Siegfried and Eisenberg (1980a, 1980b) and estimate the effect of doubleheaders on annual MLB attendance from 1920-2009. For every sub-period of our annual sample, we find the total effect of doubleheaders on season attendance to be negative. In fact for the two sub-periods 1953-1984 and 1987-2009, we find the total effects of doubleheaders on season attendance are not only negative but substantial and highly statistically significant.

Our consistent findings using both daily and annual MLB data that doubleheaders reduce season attendance leads us to question whether the widespread use of doubleheaders, at least from 1938-1984, was rational. Two of our other findings, however, provide strong evidence that baseball fans are very responsive to incentives. First,

[5] All of these studies find this simple direct effect to be positive and are roughly consistent with our findings for comparable time periods.

[6] For recent work regarding promotions, see McDonald and Rascher (2000), Boyd and Krehbiel (2006), Browning and DeBolt (2007) and Lemke et al. (2010).

from 1938-1984, doubleheaders have positive effects on attendance on the days of doubleheaders but negative effects on attendance at games surrounding doubleheaders. Second, in the post 1987 period, the substitution effects between nearby single games are sizeable and extend up to 7 days before and after each single game.

1.3 A Model of Fan Attendance Choice

We assume that fans consider their home MLB baseball schedule which consists of a series of single games and doubleheaders offered on certain dates and then choose which single games and doubleheaders to attend in order to maximize their utility over the year.[7] Our utility function offered below emphasizes two points: (1) that both single games and doubleheaders are substitutes for nearby games and (2) that doubleheaders are preferred to single games, ceteris paribus. Let the annual utility of a baseball fan be given by

$$U = \begin{array}{l} \sum_{t=1}^{T} \alpha_t(Z_t) S_t + \sum_{t=1}^{T} \left(\alpha_t(Z_t) + h\right) D_t - \sum_{l=1}^{L} \sum_{t=1}^{T} \beta_l S_t S_{t-l} \\ - \sum_{l=1}^{L} \sum_{t=1}^{T} \delta_l D_t D_{t-l} - \sum_{l=1}^{L} \gamma_l \left[\sum_{t=1}^{T} D_t S_{t-l} + \sum_{t=1}^{T} S_t D_{t-l} \right] + X \end{array} \tag{1.1}$$

In the utility function above S_t is a discrete variable equal to 1 if a single game on day t is attended and 0 if not attended. Similarly, D_t is a discrete variable equal to 1 if a doubleheader on day t is attended and 0 if not attended. X is a continuous variable measuring the amount of the numeraire good consumed per year. The function $\alpha_t(Z_t)$

[7] Of course the preseason home schedule is often modified during the season to make up for games that are rained out.

measures the direct effect on utility of attending a single game on day t. Z_t is vector of observable variables such as day of the week, month, winning percentages of the home and visiting teams, opening day, holiday, new stadium and so on that affect daily attendance. The parameter h represents the additional utility the fan gets from attending a doubleheader rather than a single game. The parameters β_l, γ_l and δ_l for $l = 1, 2, ..., L$ and $t = 1, 2, ..., T$ are all assumed to be positive.

The price of the numeraire good is normalized to 1, the price per single game and doubleheader is p and the consumers' annual income is Y. Maximization of utility subject to the budget constraint $p \sum (S_t + D_t) + X \leq Y)$ requires that the first order conditions below be satisfied:

$$\frac{\Delta U}{\Delta S_t} = \alpha_t(Z_t) - \sum_{l=1}^{L} \beta_l \left(S_{t+l} + S_{t-l}\right) - \sum_{l=1}^{L} \gamma_l \left(D_{t+l} + D_{t-l}\right) \geq p, \tag{1.2}$$

$$\frac{\Delta U}{\Delta D_t} = \alpha_t(Z_t) + h - \sum_{l=1}^{L} \delta_l \left(D_{t+l} + D_{t-l}\right) - \sum_{l=1}^{L} \gamma_l \left(S_{t+l} + S_{t-l}\right) \geq p, \tag{1.3}$$

$$\frac{\partial U}{\partial X} = 1. \tag{1.4}$$

Equation (1.2) above requires that the marginal utility of attending a single game be greater than or equal to the price of a single game. The marginal utility of attending a single game at time t is inversely related to the attendance at nearby single games up to L calendar days away. Also, the marginal utility of attending a single game at time t is inversely related to attending nearby doubleheaders up to L calendar days away. Similarly, equation (1.3) requires that the marginal utility of attending a doubleheader be greater than or equal to the price of a doubleheader. The marginal utility of a doubleheader is inversely related to attendance at nearby single games and doubleheaders up to L calendar days away.

Assuming there are always some fans on the borderline of indifference between attending or not attending either a single game or a doubleheader on any given day, attendance at single games and doubleheaders will be less when they are surrounded by nearby single games or doubleheaders. We expect the impact of a nearby doubleheader on single game attendance to be especially large because casual fans will substitute doubleheaders for nearby single games and having attending a doubleheader many fans will be much less likely to attend another nearby single game because of diminishing marginal utility.

1.4 Data and Descriptive Analysis

To investigate the presence of substitution effects in MLB, we use a collection of individual regular season game logs sorted by home team from retrosheet.org. Across seasons, the game logs are arranged chronologically starting with the 1871 season and ending with the 2009 season. Within a season, game logs record daily home team attendance and many game-specific outcomes. For example, the game logs record if the game was a night or day game, if the game was a single or a doubleheader game, and record the day of the week, month, and year the game was played. Additionally, many of the game logs offer a comprehensive list of offensive and defensive statistics for each game as well as other information.

While the game logs contain a rich account of individual home games, prior to 1938 the game logs have numerous missing values for attendance. As such, our sample on MLB daily attendance begins in 1938.[8] However, as far back as 1920, the game logs have complete records on the annual number of doubleheaders and annual games per team. Thus for analysis using only annual observations similar to Seigfried and Eisenberg (1980a, 1980b), we use the number of doubleheaders from the Retrosheet data from 1920 to 2009 along with Lahman's annual baseball attendance figures for 1920 to 2009.

When we use the word doubleheader we mean single-priced doubleheader because such an event offers fans an opportunity to see two games for the price of one. Separate-admission doubleheaders offer no such incentive; they are merely two games held on the same calendar day. To distinguish between these two types of doubleheaders in the Retrosheet game logs we had to make some simplifying assumptions and check their validity. The details of which are discussed in Appendix A. Figure 1 shows the percentage of total MLB games that are single-priced doubleheaders from 1920 to 2009 and demonstrates the importance of doubleheaders in MLB prior to 1970. At their peak in 1945 single-priced doubleheaders accounted for 48% of the total games played. Also note that in every year from 1926 to 1967, the percentage of single-priced doubleheaders accounted for at least 20% of the total games played.

[8]As a rough check on the validity of the Retrosheet daily attendance figures from 1938-2009, we aggregate the Retrosheet daily attendance figures for each year and then compare them with Lahman's annual MLB attendance figures. The results of this comparison and the frequency of missing values are discussed in Appendix A. Lastly, for more on Sean Lahman's Baseball Archive, see www.baseball1.com.

Additionally, we analyzed the distribution of doubleheaders by month and day of week from 1920 to 2009. Across months, the distribution was found to increase during the first two months of the season (4% in April to 10% in May), peak in July (17% in July), show some persistence in August and September (15% for each month) and decline in October (10% in October). For the days of the week, the distribution of doubleheaders is essentially uniform from Monday to Saturday with values ranging from 7% to 11% but then peaks at 32% on Sundays. Such patterns could exist in response to known seasonality in the price elasticity of demand for MLB in general. For example, if teams forecasted more elastic demand during times where the propensity of leisure activities is highest (i.e. during the summer vacation months and weekends), then teams may tend to schedule doubleheaders for those times in an effort to maximize revenue.

Next, we analyzed the average attendance of doubleheaders and single games over time. Specifically, Figure 2 graphs the average attendance of doubleheaders and of single games from 1938 until 2009 indicated with triangles and boxes, respectively. As shown, the average attendance for single-priced doubleheaders exceeded that of single games until the late 1980s; thereafter, the comparison becomes noisy due to the relatively small number of doubleheaders played from 1990 to 2009.

In Figure 2, we also plot twice the average attendance of single games over time, shown as the dotted line. Doing so will enable us to consider two simple measures of the attendance impact of doubleheaders, a topic we re-visit in sections 5, 6 and 7. The first simple measure we consider is the direct effect of a doubleheader, which we initially define as the difference between the average attendance of doubleheaders and the average attendance of single games. From our theory, the additional utility a fan receives from attending doubleheaders relative to single games is assumed to be positive, which implies attendance should be larger for doubleheaders relative to single games. Figure 3 plots the direct effect of doubleheaders (solid line), which is shown to be initially positive, slowly declining over time and negative for most of the 1990s and beyond.

While the direct effect of doubleheaders is an interesting marginal comparison, it fails to offer an approximation of the total effect because the numbers of games are not held constant in the calculation. As such, the second simple measure we consider is the total effect of doubleheaders, which we initially define as the average attendance of doubleheaders minus twice the average attendance of single games. This allows us to mimic a simple policy experiment: would season attendance be larger if a team played two separately priced single games or if it played one single-priced doubleheader. Using this simple measure, Figure 3 illustrates that the total effect of doubleheaders is negative for most years as shown by the dashed line. The only exception occurs approximately during the time period between 1938 and 1944.

While the two simple effects shown in Figure 3 are suggestive, they do not control for many important factors such as team effects, opening day effects, new stadium effects, winning percentages of both teams, holiday effects, and the distribution of single games and doubleheaders across days of the week and months of the year. Also, the simple total effect of doubleheaders shown in Figure 3 does not account for the possibility that doubleheaders cause fans to substitute away from nearby single games towards doubleheaders.

To test for the direct effect of doubleheaders as well as the various substitution effects implied by our model, we use regression analysis since a simple comparison of means will be unable to isolate such effects. After testing our model's predictions, we illustrate the relevance of our findings in three ways. The first is a discussion of the total effect of doubleheaders as a historical analysis. The second is a discussion of the substitution effects across single games which have contemporary relevance in determining how to schedule the various single games within a season for a given set of teams. The third is to comment on the effectiveness of promotions in increasing attendance in the presence of possible substitution effects, which again has contemporary relevance since promotions are widely used not only in professional sports but also by various retail outlets in a variety of industries.

1.5 Empirical Model and Full Sample Results

We use the Retrosheet data from 1938 to 2009 to test our theory of fan choice using the following empirical model:

$$\begin{aligned}
&Daily_{i,j,k} = \beta_1 + \beta_2 1\{D_{i,j,k}\} + \sum_{a_1=1}^{A_1} \pi_{a_1} 1\{D_{i,j,k}\} 1\{D_{i,j-a_1,k}\} \\
&+ \sum_{p_1=1}^{P_1} \pi_{p_1} 1\{D_{i,j,k}\} 1\{D_{i,j+p_1,k}\} + \sum_{a_2=1}^{A_2} \gamma_{a_2} 1\{D_{i,j,k}\} 1\{S_{i,j-a_2,k}\} \\
&+ \sum_{p_2=1}^{P_2} \gamma_{p_2} 1\{D_{i,j,k}\} 1\{S_{i,j+p_2,k}\} + \sum_{a_3=1}^{A_3} \pi_{a_3} 1\{S_{i,j,k}\} 1\{D_{i,j-a_3,k}\} \\
&+ \sum_{p_3=1}^{P_3} \pi_{p_3} 1\{S_{i,j,k}\} 1\{D_{i,j+p_3,k}\}\} + \sum_{a_4=1}^{A_4} \gamma_{a_4} 1\{S_{i,j,k}\} 1\{S_{i,j-a_4,k}\} \\
&+ \sum_{p_4=1}^{P_4} \gamma_{p_4} 1\{S_{i,j,k}\} 1\{S_{i,j+p_4,k}\} \\
&+ \mathbf{x}^T \boldsymbol{\beta} + \xi_{i,j,k},
\end{aligned} \tag{1.5}$$

where $Daily$ is the daily attendance level of home team i on day j in year k, $1\{\bullet\}$ is the indicator operator which equals 1 if the argument inside equals 1 and equals 0 otherwise, D denotes a doubleheader game if the i, j, k^{th} observation is a doubleheader, S denotes a single game and is similarly defined, the various interactions are important towards estimating the various substitution effects implied by our theory and are explained below, $\mathbf{x}^T$ is a column vector of remaining factors which influence daily attendance and may be correlated with doubleheader or single games. Specifically it is partitioned to include factors that are team specific but vary over years and factors that are team specific but vary over days and years. Lastly, $\xi_{i,j,k}$ is the error term which captures all remaining factors that affect attendance. This error term is decomposed in the following way: $\xi_{i,j,k} = \mu_i + \lambda_k + \varepsilon_{i,j,k}$, μ_i where is a team specific factor that does not vary over time, λ_k is a year-varying factor common to all teams and $\varepsilon_{i,j,k}$ is the idiosyncratic error.

The interaction terms in equation (1.5) deserve a more thorough discussion as they are used to test our theory of fan choice. The interaction terms in equation (1.5) are the same as in our utility function given by equation (1.1) except that in equation (1.5) we do not require the leads and lags to have symmetrical effects. The first four interaction terms analyze how neighboring games impact a doubleheader. Specifically, $1\{D_{i,j,k}\}1\{D_{i,j-a_1,k}\}$ indicates that a i, j, k^{th} doubleheader was played a_1 days after a neighboring doubleheader. Next, $1\{D_{i,j,k}\}1\{D_{i,j+p_1,k}\}$ indicates that a i, j, k^{th} doubleheader was played p_1 days prior to a neighboring doubleheader. Together, these interactions will help estimate how the attendance of a given doubleheader is impacted when a neighboring doubleheader game occurs a_1 days before or p_1 days after. Similarly, $1\{D_{i,j,k}\}1\{S_{i,j-a_2,k}\}$ indicates that a i, j, k^{th} doubleheader was played

a_2 days after a neighboring single game. Lastly, $1\{D_{i,j,k}\}1\{S_{i,j+p_2,k}\}$ indicates that a i, j, k^{th} doubleheader was played p_2 days prior to a neighboring single game. These latter interactions will help estimate how the attendance of a given doubleheader is impacted when a neighboring single game occurs a_2 days before or p_2 days after.[9]

The remaining four interaction terms in equation (1.5) are similarly defined and are used to examine the impact of neighboring games on single games. Specifically, $1\{S_{i,j,k}\}1\{D_{i,j-a_3,k}\}$ indicates that a i, j, k^{th} single game was played a_3 days after a neighboring doubleheader. Next, $1\{S_{i,j,k}\}1\{D_{i,j+p_3,k}\}$ indicates a single game was played p_3 days prior to a neighboring doubleheader. Together, these interactions will help estimate how the attendance of a given single game is impacted when a neighboring doubleheader game occurs a_3 days before or p_3 days after. Similarly, $1\{S_{i,j,k}\}1\{S_{i,j-a_4,k}\}$ indicates a single game was played a_4 days after a neighboring single game. Lastly, $1\{S_{i,j,k}\}1\{S_{i,j+p_4,k}\}$ indicates a single game was played p_4 days prior to a neighboring single game. These latter interactions will help estimate how the attendance of a given single game is impacted when a neighboring single game occurs a_4 days before or p_4 days after.[10]

To link the predictions of our theory to our empirical model, we first differentiate equation (1.5) with respect to doubleheader games to get the effect of doubleheaders on daily attendance:

[9] Note, the exact number of included lags and leads of the neighboring game are denoted by A_1 to A_2 and P_1 to P_2, respectively, and are allowed to be non-symmetric.

[10] Again, the exact number of included lags and leads of the neighboring game are denoted by A_3 to A_4 and P_3 to P_4, respectively, and are allowed to be non-symmetric.

$$
\begin{aligned}
\frac{\partial Daily_{i,j,k}}{\partial D_{i,j,k}} &= \beta_2 + \sum_{a_1=1}^{A_1} \pi_{a_1} 1\{D_{i,j-a_1,k}\} + \sum_{p_1=1}^{P_1} \pi_{p_1} 1\{D_{i,j+p_1,k}\} \\
&+ \sum_{a_2=1}^{A_2} \gamma_{a_2} 1\{S_{i,j-a_2,k}\} + \sum_{p_2=1}^{P_2} \gamma_{p_2} 1\{S_{i,j+p_2,k}\}.
\end{aligned} \tag{1.6}
$$

The first parameter, β_2, represents the change in attendance from a doubleheader relative to a single game after controlling for the effects of nearby doubleheaders and single games and holding all other control variables constant. We refer to this as the direct effect of a doubleheader on attendance. Because our theory suggests doubleheaders are always preferable to single games, ceteris paribus, we expect $\beta_2 > 0$. Our theory also indicates that the additional utility a fan receives from attending a doubleheader decreases if the doubleheader is nearby other doubleheaders or single games. This implies that the attendance for a given doubleheader is decreased by neighboring games. Thus, we expect π_{a_1},π_{p_1}, γ_{a_2} and γ_{p_2} to be negative.

The remaining predictions of our theory can be illustrated by differentiating equation (1.5) with respect to single games:

$$
\begin{aligned}
\frac{\partial Daily_{i,j,k}}{\partial S_{i,j,k}} &= \sum_{a_3=1}^{A_3} \pi_{a_3} 1\{D_{i,j-a_3,k}\} + \sum_{p_3=1}^{P_3} \pi_{p_3} 1\{D_{i,j+p_3,k}\} \\
&+ \sum_{a_4=1}^{A_4} \gamma_{a_4} 1\{S_{i,j-a_4,k}\} + \sum_{p_4=1}^{P_4} \gamma_{p_4} 1\{S_{i,j+p_4,k}\}.
\end{aligned} \tag{1.7}
$$

Our theory also indicates that the additional utility a fan receives from attending a single game decreases if the single game is nearby other doubleheaders or single games. This implies that the attendance for a given single game is decreased by neighboring games. Thus, we expect π_{a_3}, π_{p_3}, γ_{a_4} and γ_{p_4} to be negative.

When estimating equation (1.5) we included a number of standard controls. Specifically, we added home team winning percentage (current and lagged one year) and a new stadium dummy to control for home team factors that vary across seasons. Also, we added visiting team winning percentage (current and lagged one year), day of the week dummies, months of the year dummies, a night game dummy and a collection of holiday dummies for Opening, Memorial, Independence and Labor Day. All of which should control for relevant home team factors that may vary over days and years. To control for time-invariant home team specific effects, such as the location of a team which may persistently impact attendance, we included a set of home team dummy variables. Lastly, to control for any unobserved year effects common to all teams, such as MLB rule changes or nation-wide macroeconomic events, we included a set of year dummy variables. Almost all of these controls are highly significant and of the expected algebraic sign. They are presented and briefly discussed in Appendix A in Table 29.

As an initial specification, we test our theory empirically by estimating equation (1.5) after including all aforementioned controls and specifying that the number of lags and leads of nearby games be set to 1.[11] The results are show in column 1 of Table 1. Notice all but one of the coefficient estimates are statistically significant and all but one match in sign with respect to the predictions of our model. Specifically, the point-estimate for the variable *Doubleheader* indicates that doubleheaders relative to single games with no neighboring games within one day increase attendance by 4,102. Attendance at a given single game is lower when it is surrounded by nearby single

[11] Specifically in terms of equation (1.5), we set $L_1 = P_1 = L_2 = P_2 = L_3 = P_3 = L_4 = P_4 = 1$.

games. That is, the point-estimates for the variables *S 1 Day Prior S* and *S 1 Day After S* indicate that attendance at a given single game is lowered by 623 and 872, respectively, if there is a neighboring single game one day later or one day before.[12] Also, as one would expect, nearby doubleheader games have a more powerful effect on single game attendance. Specifically, the point-estimates for the variables *S 1 Day Prior D* and *S 1 Day After D* suggest that attendance at a given single game is lowered by 4,294 and 1,690, respectively, if there is a neighboring doubleheader one day later or one day before.

Similarly, attendance at a given doubleheader is lower when it is surrounded by nearby doubleheaders. The point-estimates for the variables *D 1 Day Prior D* and *D 1 Day After D* indicate that attendance at a given doubleheader is lowered by 4,796 and 318, respectively, when a neighboring doubleheader occurs one day later or one day before.[13] However, the latter point-estimate is statistically insignificant. Lastly, the point-estimate for the variable *D 1 Day Prior S* suggests attendance at a given doubleheader is lowered by 1,026 when a neighboring single game occurs one day later. However, the point-estimate for the variable *D 1 Day After S* indicates that attendance at a given doubleheader increases by 1,388 if a neighboring single game occurs one day prior. This latter estimate is statistically significant and conflicts with

[12] Note in column 1 of Table 1, *S 1 Day Prior S* indicates that a current single game is played 1 day prior to a neighboring single game. From equation (1.7), it follows that the point-estimate of -623 measures the loss in attendance of a given single game because it is neighbored by another single game one day later. Similar interpretations hold for *S 1 Day After S*, *S 1 Day Prior D*, *S 1 Day After D* and all similarly defined additional variables which include longer time-horizons beyond the simple case of one lag and one lead.

[13] Note in column 1 of Table 1, *D 1 Day Prior D* indicates that a current doubleheader game is played 1 day prior to a neighboring doubleheader. From equation (1.6), it follows that the point-estimate of -4,796 measures the loss in attendance of a given doubleheader because it is neighbored by another doubleheader one day later. Similar interpretations hold for *D 1 Day After D*, *D 1 Day Prior S*, *D 1 Day After S* and all similarly defined additional variables which include longer time-horizons beyond the simple case of one lag and one lead.

our theory.

For our final specification, we investigated the persistence of these nearby effects by including a variety of lags and leads and re-estimating equation (1.5). Specifically, we analyzed models that allowed the number of lags and leads to equal 7, 14, 20, 30 and 40 and used the statistical significance of the relevant point-estimates across such models as a guide for our final specification which is shown in column 2 of Table 1. Out of the 27 point-estimates in column 2 which are relevant to our theory, 23 are the correct sign and 21 are statistically significant. Doubleheaders are still shown to increase attendance by 3,458 relative to single games when there are no nearby home games. We continue to find that attendance at single games is lower when they are neighbored by other single games. This negative impact from nearby single games is fairly persistent even when these nearby single games occur seven days later or seven days before a given single game. Nearby doubleheaders for up to three days continue to decrease attendance at single games. Similarly, attendance at doubleheaders is lower if there are neighboring doubleheaders within two days. However, we continue to find that attendance at doubleheaders increase—this time by 1,365—if a neighboring single game occurs one day prior. This effect continues to conflict with our theory.

The empirical results presented in Table 1 provide strong evidence that baseball fans are very responsive to incentives and verify, with few notable exceptions, the presence of the various short run substitution effects implied by our theory. Surprisingly, the substitution effects between current single game attendance and nearby single games appear to be significant up to seven days before and after.

1.6 Stability of the Regressions over Time

In Figure 3, a comparison of simple means indicated that the simple direct effect of doubleheaders on daily attendance likely changed over time. To identify the time periods where the relationship between doubleheaders and daily attendance is stable, we employ an iterative Chow Test for structural change.[14] Specifically, we began by estimating equation (1.5) under our preferred specification for the years 1938 to 1939 but added an interaction term of doubleheader and a dummy variable for 1939. Next, we performed an F-test on the joint significance of this interaction term. If at the 10% level, the F-test failed to reject the null, then an additional year was added to the regression, an additional interaction term was added to the model and the F-test was performed again. If at the 10% level, the F-test rejected the null, then a break was found in the year that was included and the iterative procedure restarts by using data from the identified break year and the next year.[15]

This procedure detected 14 possible breaks. To illustrate these breaks, we re-estimated equation (1.5) under our preferred specification for the years where the relationship between doubleheaders and attendance was stable as indicated by the iterative Chow Test. Next, we graphed the point-estimate on doubleheaders for each set of years where the relationship was predicted to be stable. The plot is shown in Figure 4 and the flat portions motivated our selection of the following three stable sub-periods: 1938 to 1947, 1953 to 1984, and 1987 to 2009.

For each of these three sub-periods we re-estimate equation (1.5) under our preferred specification. The results are shown in Table 2. In the first sub-period between

[14] For more on the basic Chow Test for structural change, see Wooldridge (2003).

[15] Consider the following simple example when only data from 1938 to 1939 is used. If the F-test failed to reject the null, then data for 1940 was added and the process re-starts. If the F-test rejected the null, then 1939 is where a break occurs. The process re-starts using data for 1939 and 1940.

1938 and 1947 when the structure of MLB was relatively stable and doubleheaders were very common, we estimate that the direct effect of doubleheaders, β_2, is to increase attendance on the days of doubleheaders by 3,713 relative to single games. We also find that attendance of a given single game is lowered by 1,621 and 1,382, respectively, when a neighboring doubleheader occurs one day later or one day before. However neighboring doubleheaders beyond one day failed to alter attendance for a given single game in a statistically significant way. Additionally, we find that attendance of a given doubleheader is lowered by 2,207 when a neighboring doubleheader occurs one day later; but, all other doubleheader interactions are statistically insignificant. Regarding interactions between single games, we find that current single game attendance decreases by 976 when a neighboring single game occurs one day before. All other single game interactions are found to be insignificant. Finally, contrary to the expectations of our theory we find that attendance of a given doubleheader increases by 993 and 1,396, respectively, when neighboring single games occur one day later or one day before.

In the second sub-period between 1953 and 1984, we find that the simple direct effect of doubleheaders, β_2, increased attendance by 4,177 on the days of the doubleheaders. We find that attendance of a given single game is lowered by 2,569 and 1,620, respectively, when a neighboring doubleheader occurs one day later or one day before. In addition, neighboring doubleheaders beyond one day do begin to alter the attendance level of a given single game in a statistically significant way. Additionally, we find that attendance of a given doubleheader is lowered by 3,632 and 2,299, respectively, when a neighboring doubleheader occurs one day later or one day before. In fact, such neighboring effects persist when neighboring doubleheaders occur two days later or two days before. Regarding interactions between single games, we find that current single game attendance decreases by 687 when a neighboring single game occurs one day before. Additional single game interactions are found to be negative and statistically significant; however, some are positive and statistically significant. Lastly, we continue to find evidence that attendance of a given doubleheader increases at least when a neighboring single game occurs one day before. However, this increase is not statistically significant.

In the final sub-period between 1987 and 2009, when the relative frequency of doubleheaders was very small, we no longer find that doubleheaders increased attendance on the days of doubleheaders. This is consistent with our earlier discussion of Figures 2 and 3 where we found that mean attendance at single games exceeded mean attendance at doubleheaders in this last sub-period. Interestingly, however, we do find strong evidence that attendance of a given single game is reduced when neighbored by doubleheaders which occur three days later or three days before. Perhaps with the decline in the use of doubleheaders in this last sub-period, doubleheaders were used disproportionately by teams with weak average attendance and that this accounts for the negative coefficient on the doubleheader dummy variable. Yet, the correlation between the number of doubleheaders and home team season winning percentage for this sub-period was fairly low at -.18.

Alternatively, weather may also be biasing our results. In particular, doubleheaders during the last sub-period were predominately used as makeups because of rained out games. Since rain and attendance are likely to be negatively correlated and doubleheaders and rain are positively correlated our current point-estimate may be negatively biased. Retrosheet's event files record weather events as 0 for unknown, 1 for none, 2 for drizzle, 3 for showers, 4 for rain and 5 for snow.[16] However, after reviewing a number of select years and teams, we found most of the entries were set to unknown. Additionally, the Retrosheet's schedule files do contain a field that indicates which games were postponed or rescheduled but all values were missing. Nevertheless, by using the analysis of the 2007 MLB season by Lemke et al. (2010), our data and a simplified omitted variable bias rule, we found that our current point-estimate likely biases the true impact of a doubleheader downward by at most 2,393.[17] Thus, as a crude correction, our results would then indicate that the simple direct effect of doubleheaders may have been small but positive during this final sub-period.

With regard to the substitution effects between nearby single games, we find very strong and relatively persistent effects in the last sub-period. The estimated impact of neighboring single games which occur up to seven days after a given single game are all negative. Also, six out of seven point-estimates are statistically significant and the largest impact—a loss of 854—occurs when the neighboring single game is one day after a given single game. For the estimated impact of neighboring single games which occur up to seven days before a given single game, six out of seven point-estimates are negative; five of which are statistically significant. The largest

[16] We want to thank J.C. Bradbury for making us aware of this.

[17] To clarify, assume the missing variable is the indicator variable equaling one if it rained and zero otherwise. Assume that all other included regressors other than the variable Doubleheader are uncorrelated with the indicator variable for rain. Then, following Wooldridge (2003), the bias in β_2 during this last sub-period equals $\beta_{Rain}\tilde{\delta}_{DH,Rain}$, where β_{Rain} is the true impact of rain on daily home team attendance and $\tilde{\delta}_{DH,Rain}$ is the slope coefficient from the regression of the rain indicator variable on the variable Doubleheader. After some algebra, the bias equivalently equals $\beta_{Rain}\left(\sigma_{Rain}/\sigma_{DH}\right)\rho_{Rain,DH}$, where σ_{Rain} denotes the standard deviation of the indicator variable for rain, σ_{DH} is similarly defined but for the variable Doubleheader and $\rho_{Rain,DH}$ denotes the correlation between these two variables. Assume the correlation is one. Lemke et al. (2010) provide estimates of β_{Rain} and σ_{Rain} as they examined, among other factors, the impact of rain on daily attendance for the 2007 MLB regular season. Their most negative estimate for β_{Rain} is -826 and their estimate of σ_{Rain} is .144. Based on our data, our estimate for σ_{DH} for the 2007 season is .0497. Hence, the size of the bias is -2,393, which represents an upper bound estimate in absolute value.

impact—a loss of 728—occurs when the neighboring single game is three days before a given single game. In fact, summing all the interaction terms between single games in column 3 of Table 2, we estimate in the 1987-2009 sub-period that a single game that is both preceded and followed by seven single games will have its attendance reduced 5,052. Lastly, we also find, contrary to our theory, in this last sub-period as we did in the first two sub-periods that doubleheaders that follow a single game have greater attendance.

To examine the sensitivity of our results presented in Tables 1 and 2, we examined some alternative specifications. Our estimates thus far specifically assume that the year dummy variables accurately account for any unobserved year effects common to all teams. Such factors would include, for example, rule changes or scheduling changes set by the league during the off-season. Adequately controlling for unobserved scheduling changes, especially with respect to the use of doubleheaders, is clearly an important control in our analysis. Thus, we considered an alternative specification which controlled for both unobserved year effects common to all teams by including year dummy variables and unobserved year effects specific to teams by including the interaction of year and home team dummy variables. We found the point-estimates from this alternative specification to be very similar in algebraic sign and magnitude to those of our preferred specification in Tables 1 and 2.[18]

Lastly, we admit that daily attendance is censored from above by stadium capacity and as such our point-estimates may be biased. However, after collecting data on stadium capacity from a number of sources, the proportion of sellouts—games with attendance at or above recorded capacity—equalled 3.6% for our full sample and

[18]See Table 30 in Appendix A for the results.

8.9%, 1.3% and 4.3% for our three sub-periods, respectively.[19] Given so few games are actually sold out in our sample, we suspect the bias in our results is small. We did, nonetheless, examine the possibility that sellouts may be driving our finding that doubleheaders have a positive day of effect for the full sample and the 1938-1947 and 1953-1984 sub-periods. Specifically, we dropped all sellouts from our sample and re-estimated our preferred specification. Again, we found the point-estimates from this alternative specification to be very similar in algebraic sign and magnitude to those of our preferred specification in Tables 1 and 2.[20]

1.7 The Total Effect of Doubleheaders on Season Attendance for the 3 Sub-periods

Equation (1.6) gives the direct effect of a doubleheader on attendance on the day of the doubleheader. This direct effect depends on the number and type of games surrounding the doubleheader. To get the total effect of a doubleheader on season attendance we must also subtract the effect of a doubleheader on attendance at nearby games as well as the attendance lost by playing one less single game. This latter opportunity cost of doubleheaders occurs because all MLB teams were scheduled to play the same number of home games per season. We estimate the foregone attendance by playing a doubleheader in a given year to be that year's mean single game attendance.

[19] Capacity data came from a variety of sources including: baseball-almanac.com (1920-1948), the Dope Book (1949-1985), the AL Red Book and NL Green Book (1986-2009) and ballparks.com. Specifically, ballparks.com was used for the following teams and years: Cleveland Indians (1920-1948), Philadelphia Phillies (1920-1937), St. Louis Cardinals (1920-1948), St. Louis Browns (1920-1948), Arizona Diamondbacks (2009), Toronto Blue Jays (2009), Washington Nationals (2009). Also, for a complete description of all statistics contained in the Dope Book series see McConnell (1991).

[20] See Table 31 in Appendix A for the results.

For our 3 sub-periods, 1938-1947, 1953-1984 and 1987-2009 the mean single game attendance are, respectively, 9,992, 17,204 and 29,032. The mean single game attendance figures are large enough to make the total effect of doubleheaders negative for all 3 sub-periods, *without even taking into account the negative effects of doubleheaders on attendance at nearby games*. For example, in the first sub-period the total effect of a doubleheader that has no nearby games, which generates the largest possible direct effect, is simply the simple direct effect (β_2) 3,713 minus the mean single game attendance 9,992 which equals -6,279. This serves as a lower bound estimate of the total effect of doubleheaders on season attendance in absolute value because it does not take into account the negative effect of doubleheaders on nearby game attendance.

For the second sub-period, 1953-1984, the lower bound estimate of the total effect of a doubleheader in absolute value is again a doubleheader with no nearby games, the simple direct effect (β_2) 4,177 minus the mean single game attendance 17,204 which equals -13,027. For the last sub-period 1987-2009 when the relative frequency of doubleheaders were very low the lower bound estimate of the total effect of a doubleheader in absolute value is simply the simple direct effect (β_2) -2,320 minus the mean single game attendance 29,032 which equals -31,352. Clearly, in this last sub-period doubleheaders were ineffective in increasing season attendance.

For all 3 sub-periods we find the primary opportunity cost of doubleheaders, the loss in attendance from playing one less single game, exceeds any possible direct gain in attendance from playing doubleheaders. We assume the loss in attendance from playing one less single game is adequately measured by that year's mean single game attendance. Of course it is possible that mean single game attendance may exceed the actual attendance teams could achieve by playing one more single game instead of a single-priced doubleheader. However, even if the loss in attendance from playing one more game is only 50% of mean single game attendance the *largest possible* total effects of doubleheaders—doubleheaders that have no nearby games—in all three sub-periods are still found to be negative. They are, respectively, -1,283, -4,425 and -16,836.

It is difficult to escape the conclusion that the total effect of doubleheaders on season attendance has been negative over the 72 year period from 1938-2009. One might argue in some circumstances doubleheaders are necessary in order to get the required number of home games played. We admit this but presumably teams have the option of playing separate-priced doubleheaders instead of single-priced doubleheaders. In the next section, we analyze the total effect of doubleheaders on season attendance by using annual MLB data. One advantage of using annual data is that the annual data on team attendance is available back to 1920.

1.8 Results for MLB Annual Attendance Data

As mentioned earlier, Siegfried and Eisenberg (1980a, 1980b) estimate the total effect of doubleheaders on annual minor league team attendance by regressing annual team attendance on the annual number of doubleheaders, the number of home dates and other controls. They find for the period 1973-1977 that the annual number of doubleheaders does not have a significant effect on annual minor league team attendance.

We estimate a similar model using annual MLB data for various sub-periods of our full sample, 1920 to 2009, most of which correspond to the time periods analyzed in Tables 1 and 2:

$$
\begin{aligned}
&Total\ Attend_{i,t} = \theta_1 + \theta_2 \#Double_{i,t} + \theta_3 \#Home\ Games_{i,t} \\
&+\mathbf{x}^T \boldsymbol{\theta} + \mu_i + \lambda_t + \nu_{i,t},
\end{aligned}
\tag{1.8}
$$

where $Total\ Attend_{i,t}$ denotes the total attendance for team i in year t, $\#Double_{i,t}$ denotes the total number of single-priced doubleheaders for team i in year t, $\#Home\ Games_{i,t}$ denotes the total number of home games for team i in year t, $\mathbf{x}^T$ is a column vector of remaining observable factors including the season winning percentage, the lag in season winning percentage and the number of night games for home team i in year t, μ_i is a year-invariant effect specific to team i, λ_t is a year-specific effect common to all teams, and $\upsilon_{i,t}$ is the remaining disturbance term capturing all remaining factors.

The coefficient of interest in equation (1.8) θ_2 has a similar interpretation to our total effect estimate based on daily MLB data since the number of home games enters as a control. Specifically, θ_2 measures the change in total annual home team attendance by increasing the number of doubleheader games by one, holding the number of home games constant. As such, the point-estimate for θ_2 will serve as an additional robustness check of our prior results. Lastly, the estimates for equation (1.8) are shown in Table 3 for a variety of sub-periods most of which are similar to the sub-periods analyzed in Table 2.

For the full annual sample 1920-2009, the estimated total effect of an additional doubleheader was a decline in attendance by 25,479. In the period 1938-2009, which is the full sample for our daily data, the total effect rises in absolute magnitude and indicates a loss in attendance by 34,804. For every sub-sample, the estimated total effects of doubleheaders on season attendance are negative and, with the exception of the 1920-1937 period, are statistically significant as well. Consistent with the pattern seen in our previous analysis, the estimated total effects of doubleheaders on season attendance are very large and negative in the sub-periods 1938-1947, 1953-1984 and 1987-2009, where the total effects are -11,536, -38,584 and -106,338, respectively. For these latter two sub-periods, these estimates of the total effects of doubleheaders from the annual regressions are substantially more negative than the lower bound estimates in absolute value of the total effect of doubleheaders from the daily regressions.

It is interesting that the estimate of the total effect of doubleheaders in the first sub-period 1920-1937, -5,320, is the smallest in absolute value in Table 3 and is also not statistically significantly different from zero. It could be that in earlier periods of baseball history or even from 1920-1937, doubleheaders did have a positive effect on season attendance and were simply continued over time because of tradition.

1.9 Conclusion

In this paper, we developed a model of daily baseball attendance based on utility maximizing behavior that incorporates substitution effects between nearby home games. Specifically, our model made three predictions: (1) attendance at single games and doubleheaders will be less when they are surrounded by nearby single games or doubleheaders; (2) the impact of a nearby doubleheader on single game attendance should be especially large because of the enhanced effect of doubleheaders on diminishing marginal utility from attending nearby games and because fans should be substituting doubleheaders for single games; and, (3) doubleheaders should have a positive impact on attendance on the day of doubleheaders. We then tested our model using daily MLB data from 1938-2009.

In the most recent sub-period 1987-2009, we found that neighboring single games up to seven days before and seven after had a negative and statistically significant effect on the attendance level of a given single game. Also in this sub-period, we found that neighboring doubleheaders up to three days before and three days after had even stronger negative effects on the attendance level of a given single game. We believe these negative interactions between single and nearby single games and single and nearby doubleheaders provide strong evidence for diminishing marginal utility from attending nearby games. Lastly, up until 1984, doubleheaders had positive and statistically significant effects on attendance on the days of doubleheaders.

After dividing our sample into 3 sub-periods—1938-1947, 1953-1984, and 1987-2009, we found evidence that neighboring single games with lags and leads of at least one day negatively impacted the attendance level of a given single game for each sub-period. We also found in the first two sub-periods that doubleheaders had positive and statistically significant effects on attendance on the days of doubleheaders. In addition, neighboring doubleheaders had negative and statistically significant effects on attendance on given home games, either single games or doubleheaders. In the last sub-period, we found no evidence that doubleheaders increased attendance even on the days of doubleheaders.

We used our daily results to estimate lower bound estimates of the total effect of doubleheaders on season attendance in absolute value and found for all sub-periods that doubleheaders had a negative effect on season attendance. We checked these calculations by estimating the total effect of doubleheaders on attendance using annual

MLB attendance over a longer period of time from 1920-2009. For every sub-period of our annual attendance sample, we found that the total effect of doubleheaders on season attendance was negative and, except for the sub-period 1920-1937, statistically significant. For the two sub-periods 1953-1984 and 1987-2009, we found the total effects of doubleheaders on season attendance were not only negative, but substantially larger than our lower bound estimates based on daily data.

Our consistent findings using both daily and annual data that doubleheaders reduce season attendance leads us to question whether the widespread use of doubleheaders, at least from 1938-1985, was rational from the team owners' viewpoint. Note that concession, parking and merchandise sales are generally assumed to be positively correlated with attendance so unless per capita merchandise, parking and attendance sales are substantially higher during doubleheaders one cannot appeal to these other sources of revenue to justify doubleheaders. Additionally, our finding that nearby single games are substitutes, particularly in the sub-period 1987-2009, has implications for MLB scheduling. The MLB leagues should consider the negative interactions between nearby home single games in setting baseball schedules and avoid scheduling too many consecutive home games in a given ballpark.

Lastly, viewing doubleheaders as a promotion since they offer fans a two-for-one opportunity, our results that neighboring doubleheaders dampen the attendance of a given single game may serve as a warning concerning the effectiveness of other promotions. Specifically, it may be the case that traditional promotions in MLB and other sports—including fireworks shows, reduced concession prices and free souvenirs—generate substantial substitution effects between neighboring games. Future research may investigate whether substitution effects offset any day of effects from promotional giveaways. Such findings would be important to team marketers since their goal is to use promotions to increase total season attendance subject to scarcity constrains, not to simply increase attendance for home games with promotions at the expense of home games without promotions.

1.10 Tables

Table 1. Full Sample Daily Regression Results for Two Specifications.

	(1)	(2)
	One Lag and Lead	*Additional Lags and Leads*
Doubleheader	4,102***	3,458***
	(446)	(450)
S 1 Day Prior S	-623***	-509***
	(133)	(129)
S 1 Day After S	-872***	-805***
	(109)	(112)
S 1 Day Prior D	-4,294***	-4,186***
	(332)	(336)
S 1 Day After D	-1,690***	-1,683***
	(284)	(281)
D 1 Day Prior D	-4,796***	-4,743***
	(435)	(427)
D 1 Day After D	-318	-359
	(547)	(549)
D 1 Day Prior S	-1,026**	-1,001**
	(393)	(388)
D 1 Day After S	1,388***	1,365***
	(413)	(415)
S 2 Day Prior S		-147*
		(75)
S 3 Day Prior S		-132
		(83)
S 4 Day Prior S		83
		(101)
S 5 Day Prior S		-490***
		(76)
S 6 Day Prior S		-139*
		(78)
S 7 Day Prior S		-228**
		(100)
S 2 Day After S		-214**
		(85)
S 3 Day After S		-404***
		(102)

S 4 Day After S		233***
		(73)
S 5 Day After S		141*
		(76)
S 6 Day After S		-88
		(96)
S 7 Day After S		-539***
		(67)
S 2 Day Prior D		-535
		(366)
S 3 Day Prior D		-650***
		(234)
S 2 Day After D		-402
		(264)
S 3 Day After D		-641**
		(268)
D 2 Day Prior D		-879*
		(476)
D 2 Day After D		-770*
		(435)
Observations	118477	118477
R-squared	0.579	0.58

Notes: Robust standard errors in parentheses, clustered by home team; * significant at 10%; ** significant at 5%; *** significant at 1%. Standard controls discussed in the text are included; most are shown in Table A2 located in the appendix. Section 4 and footnotes 11 and 12 offer the proper interpretation of all point-estimates shown above. The dependent variable is home team daily attendance.

Table 2. Three Sub-period Specification of the Full Lags and Leads Model.

	(1)	**(2)**	**(3)**
	1938-1947	*1953-1984*	*1987-2009*
Doubleheader	3,713***	4,177***	-2,320**
	(700)	(586)	(881)
S 1 Day Prior S	-331	-221	-854***
	(221)	(176)	(123)
S 2 Day Prior S	378	13	-261***
	(222)	(144)	(64)
S 3 Day Prior S	29	-251*	-292**
	(191)	(147)	(127)
S 4 Day Prior S	-365	267**	-143
	(285)	(127)	(113)
S 5 Day Prior S	-278	-314***	-300***
	(169)	(114)	(95)
S 6 Day Prior S	54	134	-314***
	(179)	(111)	(100)
S 7 Day Prior S	147	-201*	-443***
	(291)	(117)	(123)
S 1 Day After S	-976***	-687***	-662***
	(214)	(110)	(102)
S 2 Day After S	1	-100	-220***
	(203)	(135)	(66)
S 3 Day After S	-12	-267*	-728***
	(203)	(137)	(99)
S 4 Day After S	-32	280**	206**
	(178)	(119)	(90)
S 5 Day After S	168	310***	-87
	(178)	(112)	(96)
S 6 Day After S	-61	54	-501***
	(173)	(98)	(90)
S 7 Day After S	97	-317***	-423***
	(191)	(99)	(109)
S 1 Day Prior D	-1,621***	-2,569***	-2,892***
	(429)	(404)	(458)
S 2 Day Prior D	196	-63	-2,473***
	(335)	(404)	(442)
S 3 Day Prior D	126	-737***	-1,145**
	(324)	(255)	(533)

S 1 Day After D	-1,382***	-1,620***	-2,499***
	(440)	(200)	(509)
S 2 Day After D	-281	-524**	-1,836***
	(351)	(231)	(500)
S 3 Day After D	19	-630**	-1,669***
	(240)	(246)	(583)
D 1 Day Prior D	-2,207***	-3,632***	242
	(590)	(1022)	(2969)
D 2 Day Prior D	92	-1,357**	229
	(635)	(577)	(2931)
D 1 Day After D	-34	-2,299***	4,365
	(989)	(720)	(3343)
D 2 Day After D	-470	-1,490**	-1,323
	(627)	(574)	(2138)
D 1 Day Prior S	993	-367	-553
	(700)	(463)	(866)
D 1 Day After S	1,396**	749	1,445**
	(578)	(470)	(683)
Observations	9183	49007	51607
R-squared	0.59	0.506	0.556

Notes: Robust standard errors in parentheses, clustered by home team; * significant at 10%; ** significant at 5%; *** significant at 1%. Standard controls discussed in the text are included. Section 4 and footnotes 11 and 12 offer the proper interpretation of all point-estimates shown above. The dependent variable is home team daily attendance.

Table 3. Annual Results over Various Sub-periods.

	(1)	(2)	(3)	(4)	(5)	(6)
	1920-2009	*1938-2009*	*1920-1937*	*1938-1947*	*1953-1984*	*1987-2009*
# DH	-25,479***	-34,804***	-5,320	-11,536**	-38,584***	-106,338***
	(6,885)	(8,417)	(3,900)	(4,048)	(8,563)	(26,878)
# Night Games	-4,378	-7,220	20,776***	197	-5,532	625
	(2,887)	(4,369)	(4,678)	(2,116)	(4,228)	(9,538)
# Games	22,401***	28,246***	14,158**	13,321**	42,701***	26,197**
	(7,454)	(9,640)	(6,107)	(5,332)	(10,778)	(12,767)
Season Win %	23,438***	26,540***	14,050***	16,255***	27,832***	31,260***
	(2,171)	(2,457)	(1,289)	(2,187)	(2,408)	(2,887)
Lag Season Win%	11,285***	14,123***	1,318	-2,859	8,758***	30,251***
	(2,058)	(2,372)	(1,510)	(2,234)	(2,137)	(2,720)
Obs	1922	1634	288	160	686	656
R-sq	0.818	0.786	0.839	0.884	0.738	0.698

Notes: Robust standard errors in parentheses, clustered by home team. * Significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include team and year fixed effects. The dependent variable is home team season attendance.

1.11 Figures

Figure 1. Percentage of Total MLB Games Accounted for by Doubleheaders by Year (Data 1920-2009).

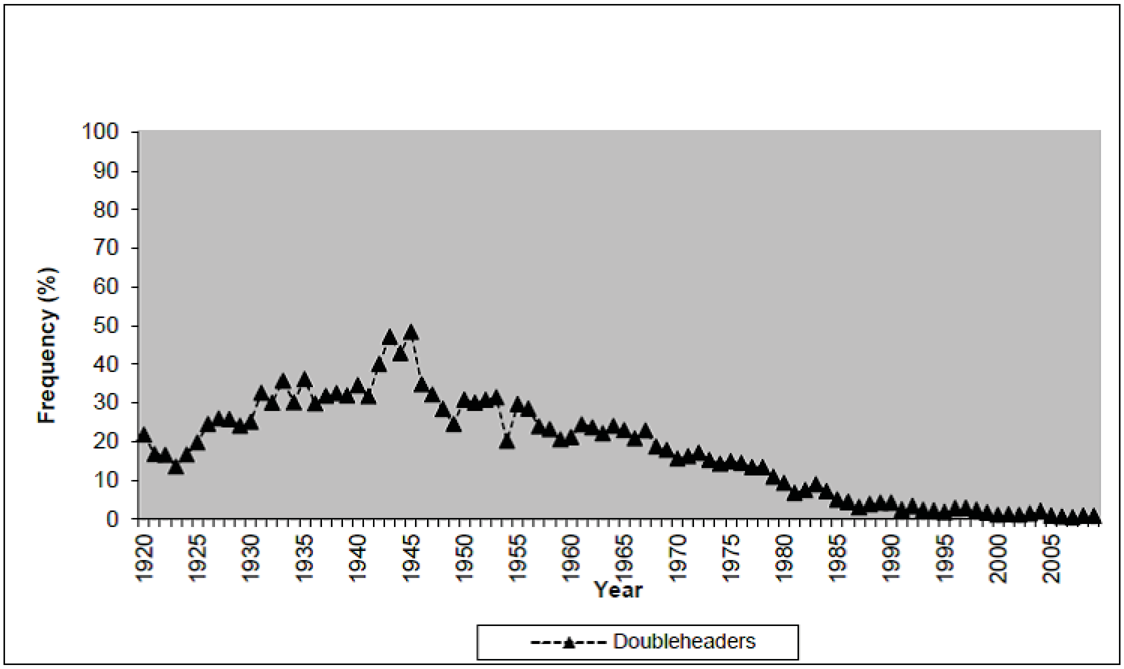


Figure 2. Average Attendance of Doubleheader and Single Games over Time (Data 1938-2009).

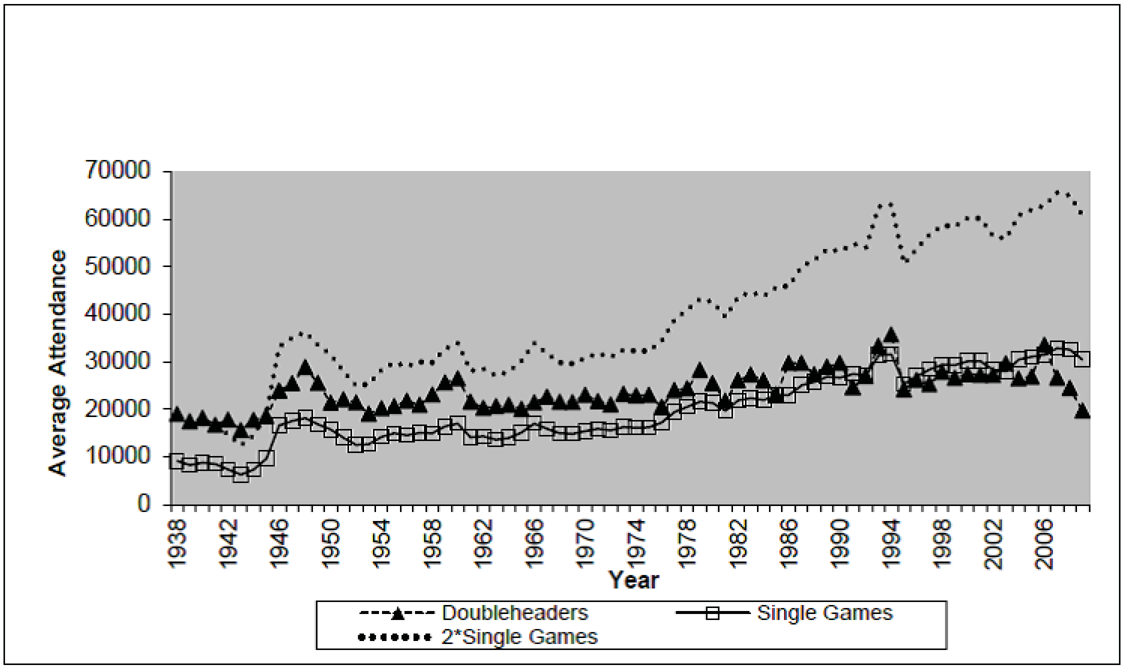


Figure 3. Illustrating Simple Direct Effects and Total Effects Estimates of Doubleheaders (Data 1938-2009).

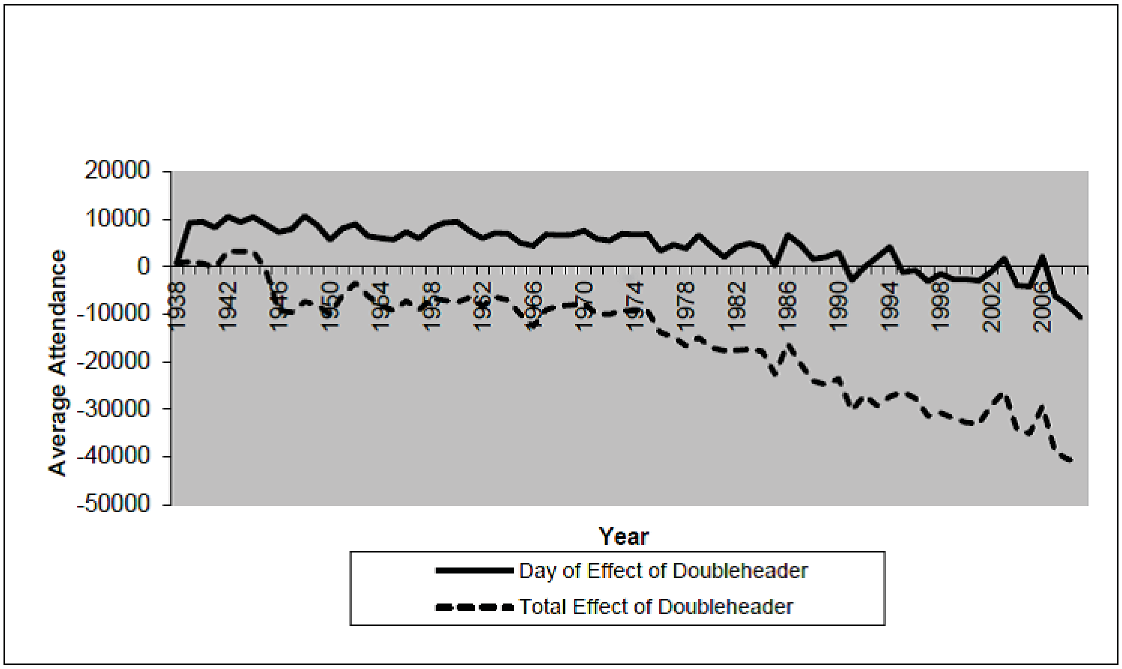
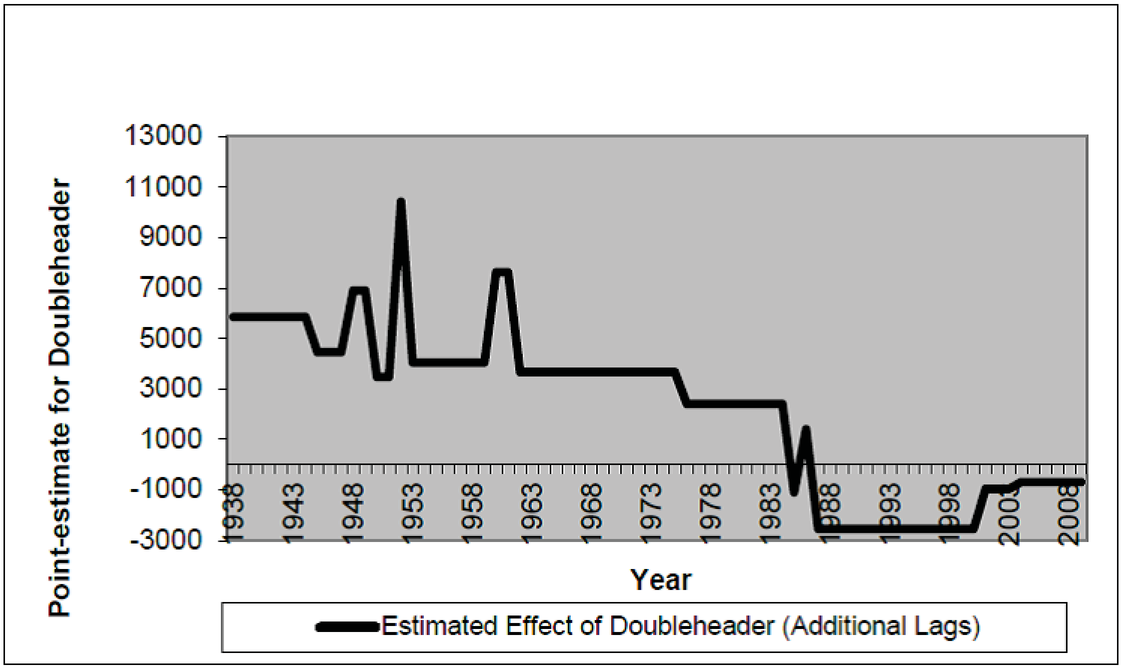


Figure 4. Estimates of Direct Effects of Doubleheader over Time.



CHAPTER II

ESSAY 2: PIGSKIN, TAILGATING AND POLLUTION: ESTIMATING THE ENVIRONMENTAL IMPACTS OF SPORTING EVENTS

2.1 Abstract

This paper estimates the environmental impact of sporting events by analyzing a collection of small typically geographically isolated cities which host at least one NCAA football team that competes in the Division I Football Bowl Subdivision (FBS) in 2010. Fixed-effects regressions controlling for differences across cities and across months suggest that cities do experience an increase in pollution levels on and around game days relative to non-game days. These marginal increases were largest in November even after controlling for weather and various trends. However, predicted levels were below EPA daily thresholds and hypothetical levels required to increase mortality rates from 0% to 1% were three to eight times larger than observed maximum game day pollution levels. Thus, the estimated marginal increases in daily pollution levels experienced by cities as a result of hosting a college football game are not hazardous and are not expected to increase mortality risks. In an effort to continue to maintain below hazardous levels of game day pollution, cities should extend their traffic-management efforts to before, during and after game days and intensify these efforts as the regular season progresses.

2.2 Introduction

Perhaps a long-held policy recommendation for increasing city economic growth has been the use of public money towards professional sports teams either by building a new stadium or by hosting a single high-profile sporting event. Typical economic justifications include the ability to attract new sources of expenditure on local goods and services, the creation of new jobs and city quality. Empirical investigations of these positive economic benefits have generally, if not universally, concluded otherwise.[1] Yet, the policy continues with recent examples including $487 million of public money towards a new ballpark for the Miami Marlins in 2012, $314.6 million of public money towards a new convention center as part of a package involving a new football stadium in Los Angeles that currently does not host a professional football team, and $498 million of public money towards a new stadium for the Minnesota Vikings in 2012.[2]

In contrast, hosting amateur sporting events with a focus towards college sports has been shown to increase economic growth in smaller more geographically isolated cities rather than large metropolitan areas. Specifically, recent evidence suggests that cities experience an increase in tax revenues and employment due to hosting college football games.[3] However, this growth has its price in that recent evidence suggests

[1]For new sources of expenditure, see Siegfried and Zimbalist (2000, 2002 and 2006), Baade (1987, 1994), Coates and Humphreys (1999) and the review by Humphreys (2009). Also, for the impact on employment and wages, see Baade and Sanderson (1997) and Coates and Humphreys (2003). For willingness to pay, see Irani (1997), Alexander et al. (2000) and Carlino and Coulson (2004, 2006) whose conclusions are contested by Coates et al. (2006). Lastly, for reviews on the economic impacts of single high-profile sporting events, see Porter (1999) and Baade (2006).

[2]See Keh (2012), Fixmer (2012) and Associated Press (2012), respectively.

[3]Coates and Depken (2009, 2010) analyzed local tax revenues and found evidence that regular season college football games generated new sources of expenditure on local goods and services. In terms of employment, Lentz and Laband (2009) found evidence of a positive relationship between college athletic revenue and employment levels in the food services industry. Indirectly related is the research by Anderson (2012) which provides evidence that college athletic success increases interest

cities experience various negative externalities ranging from public health concerns via alcohol abuse to increased crime rates.[4]

An additional and unexplored negative externality includes examining the possible environmental impacts experienced by cities when hosting college sporting events. For city officials, these impacts are important because of recent and persistent evidence suggesting that short-term changes in pollution, specifically particulate matter PM_{10} and $PM_{2.5}$[5], can increase daily mortality and cardiovascular death by .4% to 1.% (Brook et al. 2010), increase respiratory mortality by .87% to 1.68% (Anderson et al. 2012) and increase the occurrence of asthma symptoms in children by 2.3% to 2.8% (Weinmayr et al. 2010).[6] Other health effects of particulate matter include low baby birth weight and preterm birth (Shah et al. 2011) and can generally impact a very diverse at-risk population including: children, older adults, individuals with preexisting diseases—including respiratory diseases, asthma, COPD and diabetes, individuals with increased BMI and individuals of low-socioeconomic status (Sacks et al. 2011).[7]

---

in the hosting school by increasing the number of applicants, academic reputation and in-state enrollment.

[4]Specifically, Glassman et al. (2007) and Merlo et al. (2010) provide evidence of alcohol abuse and alcohol-related arrests as a result of college football home games. Further, Rees and Schnepel (2009) found that college football home games are associated with increases in assaults, vandalism, arrests for disorderly conduct and arrests for alcohol-related offenses. Lastly, Lindo et al. (2011) found team success lowered non-athlete student performance particularly among male students.

[5]According to the Environmental Protection Agency, particulate matter contain a mixture of small particles and liquid droplets where PM_{10} and $PM_{2.5}$ indicating particulate matter measuring less than 10 and 2.5 micrograms per cubic meter of air—denoted henceforth as $\mu g/m^3$ —respectively and are emitted or are formed by smokestacks, fires, power plants and automobiles. The potential health dangers derives from the small size of the particles in particulate matter which can lodge deeply into the lungs. For more information, see www.epa.gov/pm/.

[6]Note a short-term increase in pollution is measured as a 10 $\mu g/m^3$ increase in particulate matter. In terms of a time frame for exposure, Brook et al. (2010) define a short-term impact as a 10 $\mu g/m^3$ increase in particulate matter during the previous 1 to 5 days.

[7]These references are recent reviews and meta-analyses of an extensive line of literature on pollution and health. For an additional review of the literature grouped by health risks see McCubbin

Adding to the literature, this paper estimates the environmental impact of sporting events by analyzing a collection of cities which host at least one NCAA football team that competes in the Division I Football Bowl Subdivision (FBS). These cities typically have small populations relative to stadium capacity and have few alternative leisure choices that rival the attendance of college football games. Given these characteristics and the micro-level nature of the data, such a collection of cities may offer a unique quasi-experiment and thus allow causal inference. That is, given that the schedule is exogenously determined from the perspective of policy makers and residents of the city, using within-city and within-month variation of pollution on game days relative to non-game days should theoretically recover the causal effect of sporting events on pollution. However, unobserved local economic activity—by firms via unobserved production or by households via unobserved leisure activities—that in turn increases pollution may bias this approach downward and towards zero particularly if such activity systematically occurs during non-game days. Additionally, exogenous events like weather can bias this approach upwards as precipitation mitigates air pollution[8]; and, by virtue of the limited number of home games for a hosting city, precipitation would impact non-game days more so than game days.

Using daily data from the Air Quality System (AQS) database, $PM_{2.5}$ is used as the measure of pollution; and, according to the Environmental Protection Agency (EPA), it measures particles found in the air measuring less than 2.5 micrometers in diameter which come from various combustion activities including motor vehicles, power plants,

(2011) whose analysis of alternative particulate matter thresholds is summarized by the American Lung Association et al. (2011).

[8] This process is referred to as washout, see McGregor (1999) who also discusses a similar process called rainout where the formation of rain can further purge the atmosphere of gases and other suspended particles.

wood burning and certain industrial processes.[9] Supplementing the AQS data with college football data for the 2010 regular season from NCAA.org from September to November and daily weather measures from the Global Historical Climatology Network (GHCN), four questions are examined. First, do cities experience an increase in $PM_{2.5}$ on game days; that is, do cities experience a day-of effect? Second, in an effort to estimate the effects of travel, do cities also experience an increase in $PM_{2.5}$ the day before or day after game days; that is, do cities experience a commuting effect? Third, do these effects change over the course of the regular season? Fourth, do these effects pose serious health-risks?

Using city and month fixed-effects, baseline results for the simple day-of effect suggest that $PM_{2.5}$ levels increase by .72 µg/m^3 on game days relative to non-game days. Extending the analysis by estimating both commuting and day-of effects, results suggest an increase in $PM_{2.5}$ by .85, .82 and .26 µg/m^3 for days just before, on and just after game days relative to non-game days, respectively. Adding weather controls and additional controls for common or city-specific trends in pollution tended to reduce the magnitude of these point-estimates—particularly for the day-of and day-after game day effects.

For changes over the course of the regular season, the results on the commuting and day-of effects were mixed for September. $PM_{2.5}$ increased by .19 µg/m^3 on days just before a game day relative to non-game days; however, $PM_{2.5}$ decreased by .86 µg/m^3 and .91 µg/m^3 on game days and just after game days relative to non-game days, respectively. These latter estimates may suggest the presence of unobserved lo-

[9] See www.epa.gov/pmdesignations/faq.htm for an overview of particulate matter, specifically $PM_{2.5}$, by the EPA and see www.epa.gov/airquality/particlepollution/health.html for a summary regarding impact of $PM_{2.5}$ on health.

cal activity which may be more prevalent in the month of September possibly due to the start of the fall semester or by weak interest by marginal fans for FBS-FCS match-ups which overwhelmingly occur during September. For October, $PM_{2.5}$ just before, on and just after game days relative to non-game days increased by .82, 2.45 and 1.91 µg/m^3, respectively. For November, the point-estimates relative to October generally increase in magnitude; $PM_{2.5}$ just before, on and just after game days relative to non-game days increased by 1.21, 3.02 and 1.80 µg/m^3, respectively. The day-of effect is statistically significant at the 1% level and the day-after effect is statistically significant at the 5% level. After adding both weather and city-specific linear time trends, these point-estimates only changed slightly with two notable exceptions: the decrease in $PM_{2.5}$ in September for days after a game day relative to non-game days becomes statistically significant at the 5% level; and, the increase in $PM_{2.5}$ in November for days just prior to a game day relative to non-game days becomes statistically significant at the 10% level.

Expressing month-specific day-of effects in terms of percent changes, a city is predicted to experience a 24% to 31% increase in $PM_{2.5}$ when hosting a game in October or November, respectively. While these environmental impacts seem large, the predicted levels of $PM_{2.5}$ on days when a city hosts a game in October or November are well below the approximate hazardous baseline of 35 µg/m^3 established by the EPA and for all but one city are also below the recommended level of 25 µg/m^3 by the World Health Organization (WHO).[10] For example, on average, a city is predicted to experience a $PM_{2.5}$ level of 11.79 µg/m^3 and 13.66 µg/m^3 the day-of a game when it

[10] See http://www.epa.gov/air/criteria.html and WHO (2005). The 25 µg/m^3 threshold is also supported by the American Lung Association, see American Lung Association et al. (2011) and McCubbin (2011).

is played in October or November, respectively. The by-city predictions range from 6.71 to 17.48 $\mu g/m^3$ for hosted game days in October and from 7.86 to 26.96 $\mu g/m^3$ for hosted game days in November. Lastly, using short-term mortality estimates from recent publications, hypothetical game day levels of $PM_{2.5}$ required to increase the risk of cardiovascular mortality, respiratory mortality and asthma from 0% to 1% are calculated assuming air quality standards are set to month and city-specific non-game day averages of $PM_{2.5}$. These hypothetical game day pollution levels corresponding to a one percentage-point increase of cardiovascular mortality, respiratory mortality and asthma are on average 8, 5 and 3 times larger, respectively, when compared to observable maximum game day levels in October and November. Thus, cities experience a marginal change in pollution during game days in October and November but these changes are not expected to significantly or even marginally increase mortality risks.

In an effort to continue to maintain below hazardous levels of game day pollution, city-level policy implications include improving traffic-flow not only during but also near game days. These efforts should intensify as the regular season progresses. Additionally, cities could broadcast public service announcements regarding the possibility of sudden air quality changes—albeit these changes may be small in absolute magnitude and predicted levels should be below hazardous levels. To contain cost, these announcements could be sent via local media outlets with a focus towards particularly susceptible cohorts including children, older adults and those with preexisting respiratory conditions such as asthma.

The remainder of the paper is organized as follows. The identification framework is discussed in Section 2.3. Details on sample creation are discussed in Section 2.4.

Descriptive analysis of the sample is discussed in Section 2.5. Regression analysis and a discussion of the results are the topics of Section 2.6 and 2.7, respectively. Lastly, policy implications conclude the paper in Section 2.8.

2.3 Framework and Quasi-experiment

While most professional sporting events are hosted by large-populated cities, many smaller and geographically isolated cities host college sporting events. Of all the college sporting events, college football games have typically attracted the largest crowds and many are played in stadiums with capacity levels comparable to the population of the hosting city. Additionally, from the perspective of the hosting city, the football game schedule is exogenous since it determined—typically years in advance—by conference commissioners or school athletic directors.[11] Given these characteristics, cities which host well-attended college football games are particularly well-suited for analyzing the environmental impact of sporting events as game days and non-game days essentially create a "treatment" and "control" group for each city.

That is, when a game day occurs, it typically attracts a large mass of people to the hosting city. Given that the hosting city is small, this large mass of people may considerably increase pollution levels because of traffic, congestion and additional game-related activities such as tailgating. Thus, so long as this influx of people is predominately caused by the game, air quality measures during game days should serve as the "treatment" group.

[11] There are many examples. As of April 2011, the conference office for the Big 10 scheduled conference games for all teams in the Big 10 for the 2011, 2012, 2013 and 2014 seasons; see the press release at http://www.bigten.org/sports/m-footbl/spec-rel/040611aaa.html. Also in 2011, Northwestern and Notre Dame agreed to alternating home-home non-conference games for the 2014 and 2018 season; see http://sports.espn.go.com/chicago/ncf/news/story?id=6352027. Lastly, as of 2008, Penn State scheduled alternating home-home non-conference games against the Alabama for the 2010 and 2011 seasons and Virginia for the 2012 and 2013 seasons; see http://live.psu.edu/story/31377.

In contrast, during non-game days, the residents of the hosting city engage in normal activities and air quality returns to normal. As such, measures of air quality during non-game days should serve as the "control" group. Thus, the environmental impact of sporting events is measured by the difference between air quality measures during game days and non-game days. However, unobserved local economic activity—by firms via unobserved production or by households via unobserved leisure activities—that in turn increases pollution may bias this approach downward and towards zero particularly if such activity systematically occurs during non-game days.

For example, a new manufacturing plant could open mid-season and systematically pollute more during weekdays relative to weekends. Since game days occur predominately on Saturdays, this pollution if unaccounted for would tend to artificially increase pollution during non-game days. When compared against pollution during game days, the difference as a measure of the environmental impact would be artificially biased downward and towards zero. However, no bias would occur if the unobserved firm activity occurs prior to the start of the season. Given home games occur in a three month interval—September to November—and since economic activity within small isolated hosting cities may reasonably coincide with the beginning of the school year which predates the start of the season, the potential bias from unobserved firm production is expected to be small.

In terms of unobserved leisure activities by households, examples would include a concert by a popular performer that occurs during a non-game day and produces just as much pollution as a home football game. Also, commuting to cities before and after game days would produce pollution on non-game days. Again by similar reasoning, these unobserved leisure activities would bias the difference of pollution on game days versus non-game days downward and towards zero.

Lastly, exogenous events like weather can present an upward bias in this approach. Specifically, precipitation mitigates air pollution[12]; and by virtue of the limited number of home games for a hosting city, precipitation would impact non-game days more so than game days. If left unaccounted, this would artificially decrease pollution on non-game days more so than on game days. Thus, the difference as an estimate of the environmental impact would be biased upward.

The next section discusses how the sample was created in order to address and minimize these possible confounding factors—particularly unobserved firm production and household leisure activities. To address weather, the sample was supplemented with additional data including temperature and precipitation. Lastly, the specific sporting events were football games from the 2010 NCAA regular season.

2.4 Sample Construction

To estimate the environmental impact of sporting events, a sample was created in an attempt to minimize the possible confounding effects from unobserved local economic activity that may in turn increase pollution levels—either by firms via unobserved production or by households via unobserved leisure activities—particularly during non-event days. In constructing the sample, a city was included so long as the following three conditions held:

[12] See McGregor (1999).

(1) The city does not host a major professional sports team.[13]

(2) The city hosts at least one FBS team.[14]

(3) The city has a common measure of pollution at the day-level for at least some game and non-game days from September 2010 to November 2010 with other cities in the sample.

Condition 1 rules out the largest potential cities which due to their large populations have the greatest likelihood of unobserved local economic activities. Condition 2 ensures that the city hosts at least one team from the highest quality subdivision of college football but could host additional teams from other relatively lower quality subdivisions. In turn, this restriction increases the likelihood of analyzing games played in the largest stadiums and as a result are the most-followed and attended. Lastly, condition 3 is simply a feasibility requirement; a common pollution measure is required at the day-level in order to estimate the environmental impacts of college football games.

For specific pollution measures, daily averages of hourly measures of various pollutants are available via the Air Quality System (AQS) database by the Environmental Protection Agency (EPA) at the city, county and state level. In particular, the AQS provides daily summaries which track six common air pollutants—referred to as criteria pollutants—for the purposes of air quality standards set forth by the EPA under the Clean Air Act. These common air pollutants are ozone, particulate matter, car-

[13] That is, the city does not host teams that play in the National Hockey League (NHL), the National Football League (NFL), the National Basketball Association (NBA) or in Major League Baseball (MLB).

[14] Meaning, the city must host at least one FBS team but may host other college teams which compete in the remaining NCAA subdivisions including Division I Football Championship Subdivision (FCS), Division II, Division III or the National Association of Intercollegiate Athletics (NAIA).

bon monoxide, nitrogen oxides, sulfur dioxide and lead. For reasons discussed below, the air pollutant used to measure pollution is $PM_{2.5}$. According to the EPA, $PM_{2.5}$ measures particles found in the air measuring less than 2.5 micrometers in diameter which come from various combustion activities, including motor vehicles, power plants, wood burning and certain industrial processes.[15] Additionally, the EPA states that such small particles can lodge deeply into the lungs and thus are believed to pose the greatest health risks.[16] Given the potential of traffic and overall congestion associated with traveling to and from games as well as other game-day activities such as tailgating, $PM_{2.5}$ is an appropriate measure for analyzing the environmental impact of college football games.

Of the approximate 120 cities which hosted all of the FBS college teams in 2010, 33 cities satisfied conditions 1-3. The most restrictive condition was the availability of a common pollution measure at the day-level. Specifically, in 2010, there were approximately 733 college football teams.[17] Of those 733, data from the AQS on criteria pollutants were available for 296 teams. Of these 296 teams, the most common criteria pollutant was $PM_{2.5}$ at 82%. Of the 244 with the most common criteria pollutant, 61 teams belonged to the FBS. Of these 61 teams, 18 played in a city with a professional sports team and 10 played in a city with missing $PM_{2.5}$ values for all home games. With the final 33 cities, the remaining data supplements included: weather variables—minimum and maximum temperature and precipitation—at the day-level from the Global Historical Climatology Network (GHCN), football schedule

[15] See www.epa.gov/pmdesignations/faq.htm.

[16] See www.epa.gov/airquality/particlepollution/health.html.

[17] See http://prwolfe.bol.ucla.edu/cfootball/LinksList.html, maintained by Peter Wolfe and Ross Baker. Peter Wolfe also contributes to the computer rankings portion of the Bowl Championship Series (BCS) rankings further discussed momentarily and in footnote 19.

and box score data from NCAA.org and stadium latitude and longitude data from Google Maps.

Table 4 provides a list of the 33 cities in the sample. In addition, it provides basic demographic information for the year 2010 on macroeconomic factors, employment composition and average time spent commuting to work. Across all cities in the sample, the average city population level is 240,000 with a range of 42,000 for State College, PA and 790,000 for Austin, TX. Average real per capita income is $22,000 with a range from $13,000 for State College, PA and $30,000 for Austin, TX and Madison, WI. The average unemployment rate is 8.6% with a large degree of variability across cities with the lowest rate rate of 5.1% for Iowa City, IA and the highest rate of 15.6% for Kalamazoo, MI.

The second column describes the type of employment for each city and the average across all cities for five broad employment categories: management, business, science and arts occupations; service occupations; sales and office occupations; natural resources, construction and maintenance occupations; and production, transportation and material moving occupations. Overall, the employment composition is concentrated in the management, service and sales categories for each city with an overall-average of 37%, 20% and 25% of the civilian employed population age 16 and over, respectively. The remaining employment categories of construction and transportation account for the remainder with sample averages of 7% and 10% of employment, respectively. The third column lists the mean travel time to work in minutes by city. The sample average commute time is 20 minutes and few cities deviate substantially from this average.

While unobserved local economic activity, particularly by unobserved firm production, is a possibility, Table 4 provides some evidence that such activities are unlikely given that the employment categories most associated with pollution—construction and transportation—range only from a collective 9% to 25% of city employment. Additionally, the relatively high and likely stagnant unemployment rate for many of the cities in the sample may also assuage these concerns, particularly considering the time frame is only from September to November of 2010.

Similarly, in terms of unobserved local economic activity by households, larger more populated cities with high per capita income levels may in turn have more leisure choices.[18] For example, some events, like a concert performance, may generate sufficient attendance as to mimic the attendance of local football games. In turn, these events may increase pollution levels for the same reasons as football games. Table 4 may provide some evidence that these situations are possibly rare and perhaps isolated to just a few cities in the sample including Austin, TX, Columbus, OH and Louisville, KY. Lastly, the similarities in the average time commuting to work across cities may suggest similarities in the average time spent commuting to leisure activities.

Table 5 lists each FBS team hosted by the 33 cities in the sample. Provided are overall team quality measures including season winning percentage, conference winning percentage and average home game attendance for the 2010-2011 season. Of the 33 teams, 21 were bowl eligible; 3 of which played in Bowl Championship Series (BCS) bowls—Ohio State, Wisconsin and Oregon; and 1 of which played in the BCS National Championship Game—Oregon.[19] Across all teams in the sample,

[18]This logic partly borrows from the early empirical work on population thresholds—population levels required to attract additional firms in a given industry—and market structure by Bresnahan and Reiss (1990, 1991).

[19]Some institutional background is in order. Bowl games are the FBS's version of post-season games. They offer teams cash pay-outs and program exposure which aids in recruiting players. It

the average winning percentages for season and conference play is around 50% and average attendance is about 54,600. By team and city, the lowest average home game attendance was Ball State in Muncie, IN with 9,900 and the largest was Ohio State in Columbus, OH with 106,000. Of the 20 teams who play in the largest stadiums by capacity, 10 are included in the sample and are indicated with stars in Table 5.[20] Thus, the sample contains some of the most well-followed and attended teams in college football.

To summarize the location of the pollution and weather monitors, Table 6 lists by city the number of pollution monitors, the average distance between pollution monitors and a city's FBS stadium, and the average distance between weather monitors and a city's FBS stadium. The number of pollution monitors ranged from 1 to 4, with a sample average of about 2, and were typically close to a city's FBS stadium—less than 5 miles away for almost all cities in the sample. The weather monitors, while a bit farther away as most were located at a city's international or regional airport, were still less than 10 miles away for almost all cities in the sample. Thus, the monitors are a close distance to a city's FBS stadium thereby making it possible to estimate the environmental impact of hosting home games.

---

takes at least 6 regular season wins to be bowl eligible with the most prestigious being the BCS bowls. In 2010, the BCS bowls were largely determined by conference championships—teams with the best conference winning percentage at the end of the regular season or victors of a conference championship game played at the end of the regular season in December—from 6 automatic qualifying conferences: the Southeastern Conference, the Pacific 10 Conference, the Big 12 Conference, the Big 10 Conference, the Big East Conference and the Atlantic Coast Conference. A championship is determined by the BCS National Championship Game which features the number 1 BCS-ranked team against the number 2 BCS-ranked team where the BCS-rankings begin in October and conclude in December. For more information on the BCS, see bcsfootball.org.

[20] For the exact 2010 stadium rankings by capacity, see http://voices.yahoo.com/largest-college-football-stadiums-top-20-6938052.html.

The next section analyzes the sample using descriptive analysis and five primary areas are examined. First, the potential health consequences from exposure to increased levels of $PM_{2.5}$ on game days are discussed. Second, possible changes in pollution levels on game days versus non-game days are explored by analyzing differences in average $PM_{2.5}$. Third, possible trends are identified by analyzing how this differences in average $PM_{2.5}$ changes over the course of the season. Fourth, explanations for this change over time are discussed and are related to possible changes in attendance and weather conditions. Fifth and finally, descriptive statistics are averaged across cities and to the day-level in order to examine possible co-movements between attendance and $PM_{2.5}$ as evidence that cities experience both a commuting and day-of effect when hosting football games.

2.5 Descriptive Analysis

To establish the possibility that pollution levels are large enough to cause residents of game-hosting cities health problems, Table 7 lists the maximum level of $PM_{2.5}$ on game days and on non-game days by city. To better interpret the maximum values provided in Table 7 in relation to possible health concerns, the EPA establishes air quality standards for all criteria pollutants as required by the Clean Air Act. In 2010 for $PM_{2.5}$, the 24-hour standard was that a three-year average of annual 98^{th} percentile 24-hour average values recorded at each monitor site must be less than 35 micrograms per cubic meter of air.[21] This standard of 35 $\mu g/m^3$ will be considered as an approximate baseline. Comparing columns 2 and 4 of Table 7, many of the

[21] See http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_index.html.

maximum values across non-game days are larger than the corresponding maximum values across game days. However, upon further examination of the data, 12 of the 33 maximum values during non-game days were no more than 3 days away from a home FBS football game and are indicated with stars. With this in mind, two cities—Louisville, KY and Fresno, CA—experience pollution levels on or around game days in excess of 35 µg/m^3 and two additional cities—Salt Lake City, UT and Birmingham, AL—experience pollution levels close to 35 µg/m^3. As a result, there may be harmful health effects as a result of increased air pollution on game days for at least some cities in the sample.

In addition to exceeding a certain threshold level of pollution, sudden changes in air quality in a short-time interval have been shown to have various health consequences.[22] Table 8 shows the difference in average $PM_{2.5}$ for game days versus non-game days both for the entire sample and by city. Columns 3-5 provide p-values for hypothesis tests which examine if this difference is less than, different from or greater than zero, respectively. For the sample, the difference provides some evidence of a small increase in $PM_{2.5}$ of about .79 µg/m^3 on game days relative to non-game days with two-tailed and right-tailed p-values of .11 and .06, respectively.

By city, the results are mixed as 11 cities experience a marginal decrease in pollution ranging from .07 to 3.03 µg/m^3; 3 cities experience essentially no change in pollution; and, 16 cities experience a marginal increase in pollution ranging from .31 to 7.78 µg/m^3.[23] Of the 11 counter-intuitive cases where the difference in average pollution levels decreased which suggests that hosting games lowers pollution, recall

[22]For summaries and introductions to this extensive literature, see Brook et al. (2010), Weinmayr et al. (2010), Shah et al. (2011), Sacks et al. (2011), McCubbin (2011) and Anderson et al. (2012).

[23]Note, two cities—Lubbock, TX and Monroe, LA—were dropped as each only had one game day with positive $PM_{2.5}$ data.

from Table 7 that 7 of these cities had maximum $PM_{2.5}$ levels on non-game days that were within at most 3 days of a game. As a result, these values would inflate the average taken across non-game days. The non-game day averages also contain all days just before and after a home game which may capture effects from commuting to games. As a result, these days may also increase the non-game day $PM_{2.5}$ average. Thus, both sources would tend to bias the difference in means downward.

To analyze basic trends over time and to identify possible month-specific unobserved effects, Table 9 analyzes how differences in the average game day and non-game day levels of $PM_{2.5}$ differ across months, population and both months and population. Differences across population sizes were used as a parsimonious way to illustrate how trends over time likely differ by city. Small cities were defined as having population levels at or below the sample median; whereas, large cities had population levels that exceed the sample median. The difference in means is shown in column 1 and columns 3-5 provide p-values testing if this difference is less than, different from or greater than zero, respectively.

For the sample, the differences in $PM_{2.5}$ over game and non-game days exhibits a clear trend. Statistically, pollution levels are no different on game days versus non-game days in September but do increase in October and more so in November. These respective differences equal -.77, 1.69 and 2.11 $\mu g/m^3$—the former statistically indistinguishable from zero and the remaining statistically different from and exceeding zero at approximately the 5% level. By population, both small and large cities exhibit a similar qualitative trend; however, the magnitudes tend to be larger for more populated cities with statistically significant results for October and November.

To determine if these trends in pollution over time are due to increases in attendance or changes in weather conditions, Table 10 presents descriptive statistics on average game attendance, percentage of away games, temperature and precipitation by month. While attendance is not surging over time, it does remain fairly stable. There is also no descriptive evidence that attendance rapidly declines as temperature declines or rainfall increases. However, increased rainfall can decrease pollution levels.[24] In column 4, the average precipitation levels by month do increase slightly but the units of measure are in millimeters so the magnitudes are small. Nevertheless, the absence of rainfall on game days coupled with slight increases in rainfall on non-game days could potentially explain the increase over time in game versus non-game day average pollution levels.

Lastly, Figure 5 further investigates these trends by examining possible co-movements in attendance and weather conditions and in attendance and $PM_{2.5}$ at the day level over time. Specifically, averages of attendance, $PM_{2.5}$, precipitation and temperature were taken across all cities in the sample for each day starting in September and ending in November. Figure 5(a) provides evidence that attendance is fairly insensitive to changes in weather as attendance is stable to trending upward while temperature trends downward and precipitation is fairly flat with some isolated peaks. Figure 5(b) re-scales attendance and provides evidence of considerable co-movements with $PM_{2.5}$. Specifically, many of the increases and peaks in attendance correspond to similar trends in $PM_{2.5}$. In terms of the frequency of this relationship across months, these episodes tend to appear during mid September, early to late October and early to mid to late November.

[24] See McGregor (1999).

In summary, the descriptive statistics provides some evidence that pollution levels increase beyond the levels used for air quality standards by the EPA for at least some of the cities in the sample during game days. There is evidence of a marginal increase in pollution when comparing average $PM_{2.5}$ levels across game-days and non-game days. However, some of the highest levels of $PM_{2.5}$ occurred on non-game days but were close to game days. These days may involve commuting to and from games and as a result should be analyzed as game-day related. Descriptive results also suggest that the increase in pollution from game days also changes over time with a general increase over the course of the season. These changes by month may be attributable to game days via steady to increasing attendance levels or due to trends in weather via rainfall occurring more on non-game days relative to game days. Lastly, these changes by month may indicate the presence of unobserved local activity which may be more prevalent in the month of September.

In the next section, the change in daily pollution as a result of hosting a football game is estimated using regression techniques. Initially a baseline fixed-effects method is used in order to control for differences across cities and differences across months. Next, additional controls are added; these include controls for weather—temperature and precipitation—and controls for various trends in the level of pollution—trends common to all cities because of macroeconomic factors or trends specific to each city. Also considered are additional variables designed to estimate possible commuting effects before and after game days. Lastly, month interactions are added to examine how pollution changes on or around game days differ by month.

2.6 Regression Analysis

Panel data regression techniques are used in order to estimate the environmental impact of college football games with a focus on three overall approaches. Initially, the day-of effect is estimated and is defined as the increase in pollution when a city hosts a game relative to had it not hosted a game. Additional variables are added to estimate both the commuting and day-of effects where the commuting effect is defined as the increase in pollution on days just before or after a city hosts a game relative to had it not hosted a game. Lastly, month interactions are added to examine if the commuting and day-of effects change over the course of the season.

Econometrically, city and month fixed-effects are included. As a result, the day-of effect is estimated using within city and within month variation thereby comparing differences in pollution on game days to non-game days after controlling for differences across cities and months. The commuting effect is estimated by comparing differences in pollution just before or after game days to non-game days after controlling for differences across cities and months. Month-specific commuting and day-of effects are estimated by month-specific differences in pollution just before, on or after game days relative to non-game days after controlling for differences across cities and months.

Additional controls are considered. Initially, controls for weather—temperature and precipitation—are included with an expectation that the estimated day-of effect will correspondingly decline in magnitude. Next, a linear time trend is added to control for possible trends in pollution caused by common macroeconomic factors. Finally, linear time trends are allowed to be city-specific thereby controlling for possible city-specific trends in pollution caused by local unobserved firm production or unobserved leisure activities by households. Given the time horizon is only three months, the likely impact from the inclusion of these time trends on either the commuting effect or day-of effect is expected to be small.

As a baseline, the day-of effect is estimated using the following econometric model:

$$y_{ijk} = \alpha + \beta_1 HostGame_{ijk} + \mu_i + \lambda_k + \epsilon_{ijk} \tag{2.1}$$

where y_{ijk} is the level of $PM_{2.5}$ for city i on day j during month k; $HostGame_{ijk}$ is a dummy variable equal to 1 if city i hosts a football game on day j during month k; μ_i is a day-month invariant city-specific fixed effect; λ_k is a day-city invariant month-specific fixed effect; and, ϵ_{ijk} contains all remaining factors which influence $PM_{2.5}$. The parameter of interest, β_1, measures the day-of effect.

Fixed-effects estimates of the day-of effect are provided in column 1 of Table 11 and suggest that $PM_{2.5}$ levels increase by .72 µg/m^3 on game days relative to non-game days. This point-estimate is almost identical to the full sample difference in means of $PM_{2.5}$ on game days versus non-game days. However, the magnitude of this point-estimate declined by almost half after controlling for weather—shown in column 2—and adding additional controls for trend had little impact on its magnitude or statistical significance—shown in columns 3 and 4.

To estimate both the commuting and day-of effects, the following econometric model is used:

$$y_{ijk} = \begin{aligned} &\alpha + \beta_0 DayBefore_{ijk} + \beta_1 HostGame_{ijk} + \beta_2 DayAfter_{ijk} \\ &+\mu_i + \lambda_k + \epsilon_{ijk} \end{aligned} \tag{2.2}$$

where $DayBefore_{ijk}$ is a dummy variable equal to 1 for city i if day j during month k is one day before a game day and zero otherwise; $DayAfter_{ijk}$ is similarly defined but for one day after a game day; and, all remaining variables are as defined in equation (2.1). The parameters, β_0 and β_2, are designed to measure possible commuting effects. Lastly, the parameter, β_1, continues to measure the day-of effect.

Fixed-effects estimates of the commuting and day-of effects are shown in column 1 of Table 12 and results suggest an increase in $PM_{2.5}$ by .85, .82 and .26 µg/m^3 for days just before, on and just after game days relative to non-game days, respectively. Adding weather controls and additional controls for common or city-specific trends tend to reduce the magnitude of these point-estimates—particularly for the day-of and day-after game day effects.

To allow the commuting and day-of effects to be month-specific, the following econometric model is used:

$$y_{ijk} = \begin{array}{l} \alpha + \beta_0 DayBefore_{ijk} + \beta_1 HostGame_{ijk} + \beta_2 DayAfter_{ijk} \\ +\beta_3 DayBefore_{ijk} \times Oct._{ijk} + \beta_4 HostGame_{ijk} \times Oct._{ijk} \\ +\beta_5 DayAfter_{ijk} \times Oct._{ijk} + \beta_6 DayBefore_{ijk} \times Nov._{ijk} \\ +\beta_7 HostGame_{ijk} \times Nov._{ijk} + \beta_8 DayAfter_{ijk} \times Nov._{ijk} \\ +\mu_i + \lambda_k + \epsilon_{ijk} \end{array} \tag{2.3}$$

where $Oct._{ijk}$ equals 1 for all days in October; similarly, $Nov._{ijk}$ equals 1 for all days in November; the base month is September; and, all remaining variables are as defined in equation (2.2). With the addition of interaction terms, the parameters of interest are month-specific. Specifically, β_0 and β_2 measure the commuting effects and β_1

measures the day-of effect of hosting a football game in September. Similarly, β_3 and β_5 measure the commuting effects and β_4 measures the day-of effect of hosting a football game in October. Lastly, β_6 and β_8 measure the commuting effects and β_7 measures the day-of effect of hosting a football game in November.

Fixed effects estimates of month-specific commuting and day-of effects are shown in column 1 of Table 13. The results on the commuting and day-of effects of hosting a game were mixed for the month of September. $PM_{2.5}$ increased by .19 µg/m^3 on days just before a game day relative to non-game days; however, $PM_{2.5}$ decreased by .86 µg/m^3 and .91 µg/m^3 on and just after game days relative to non-game days, respectively. For October, $PM_{2.5}$ just before, on and just after game days relative to non-game days increased by .82, 2.45 and 1.91 µg/m^3, respectively. For November, the point-estimates relative to October generally increase in magnitude; $PM_{2.5}$ just before, on and just after game days relative to non-game days increased by 1.21, 3.02 and 1.80 µg/m^3, respectively. For November, the day-of effect is statistically significant at the 1% level and the day-after effect is statistically significant at the 5% level.

In general, all month-specific point-estimates increased in absolute value after adding additional weather controls, shown in column 2. Further changes were slight after adding both weather and city-specific linear time trends—column 4—but with two notable exceptions: the decrease in $PM_{2.5}$ in September for days after a game day relative to non-game days becomes statistically significant at the 5% level; and, the increase in $PM_{2.5}$ in November for days just prior to a game day relative to non-game days becomes statistically significant at the 10% level.

2.7 Discussion

The regression results have some notable similarities with the descriptive analysis. Specifically, the baseline fixed-effects estimate of .72 µg/m^3 for the day-of effect which is approximately equal to the simple difference of average $PM_{2.5}$ on game days versus non-game days at .79 µg/m^3. However after controlling for weather, the day-of estimate decreased dramatically to .44 µg/m^3 and remained approximately unchanged after additional controls for trend. The by-month regression results were also similar to the by-month descriptive statistics. The fixed-effects estimates of -.86, 2.45 and 3.02 µg/m^3 for games held in September, October and November for the day-of effect were only slightly larger in absolute value to the difference of average $PM_{2.5}$ levels on game days versus non-game days of -.77, 1.69 and 2.11 µg/m^3. After controlling for weather, the month-specific day-of effects increased in absolute value but dampened after adding additional controls for overall and city-specific trends. Lastly, the regression results indicate some evidence of a commuting-effect particularly for the months of October and November with increases in $PM_{2.5}$ ranging from .58 to 1.58 µg/m^3 for October and 2.10 and 1.92 µg/m^3 for November. However, the October point-estimates of both the commuting and day-of effects are statistically insignificant at conventional levels; whereas for November, these effects are statistically significant at the 10% level or better.

The counter-intuitive result that the commuting and day-of effects for September are negative—thereby implying hosting a game improves air-quality—is likely caused by two factors. First, non-game day pollution levels are likely elevated in September due to unobserved local activity associated with a the start of the fall semester and Labor Day weekend. Second, the following by marginal fans, particularly those who visit hosting cities to experience game-day leisure opportunities but do not attend the game, may be particularly low in September.

This drop in interest by marginal fans could occur because FBS-FCS match-ups, which historically offer the FBS team an almost assured victory, are predominantly played in September and are always a home game for the FBS team.[25] Interest by marginal fans may intensify over the course of the season as the quality of leisure opportunities increases in response to higher-quality conference match-ups which typically begin in October, intensify in November and conclude with a conference championship game in December[26]—typically played at a neutral site—or because of themed home games like "Homecoming[27]" or "Rivalry Week[28]" which typically occur in October and November, respectively. Thus, low interest by marginal fans in September combined with unobserved local activity associated with the start of the fall semester would bias the commuting and day-of effects to zero and if large enough beyond zero.[29] To relate the regression estimates to possible health concerns, Table

[25]Discussed in greater detail by Daughters (2010), of the 90 total FBS-FCS games 78 were played by September 28, 2010 featuring an FBS record of 72 wins to 6 losses. In previous seasons, the FBS record was 82-2 (2008) and 52-2 (2005).

[26]Recall from footnote 19 that for 6 conferences in 2010, conference winning percentage determines either a conference championship or the opportunity to play in a conference championship game; and in turn, conference champions receive automatic bids to BCS bowl games. This mechanism increases the importance of conference games and thus may in turn increase the quality of leisure opportunities in hosting cities during conference play which in turn attracts more marginal fans.

[27]Homecomings feature a variety of additional leisure opportunities including parades, pep rallies and other school-specific traditions.

[28]Rivalry Week ends the regular season with games starting after Thanksgiving and feature many intrastate rivalries—for example, Alabama versus Auburn, Ole Miss versus Mississippi State, Florida versus Florida State and so on—or historic rivals like USC versus Notre Dame and Ohio State versus Michigan.

[29]Interestingly, it was announced in February of 2013 that the Big 10 would stop scheduling FCS games, see http://espn.go.com/college-football/story/_/id/8942451/barry-alvarez-says-big-ten-schedule-fcs-teams.

14 shows the percent change in $PM_{2.5}$ and predicted levels of $PM_{2.5}$ when cities host a game in October and November. On average, a city is predicted to experience a 24% increase in $PM_{2.5}$ when hosting a game in October and ranges by city from 15% to 52% for Lafayette, LA and Tucson, AZ, respectively. Similarly for November games, $PM_{2.5}$ is predicted to increase by 31% with a range from 14% to 71% for Fresno, CA and Albuquerque, NM, respectively. While on a percent change basis the environmental impacts seem large, the predicted levels of $PM_{2.5}$ on days when a city hosts a game in October or November are well-below the approximate hazardous baseline of 35 $\mu g/m^3$. For example, on average, a city is predicted to experience a $PM_{2.5}$ level of 11.79 $\mu g/m^3$ and 13.66 $\mu g/m^3$ when hosting games in October and November, respectively. The by-city predictions range from 6.71 to 17.48 $\mu g/m^3$ for game days in October and from 7.86 to 26.96 $\mu g/m^3$ for game days in November. Only one city—Fresno, CA—is predicted to experience pollution levels for games played in November in excess of an alternative baseline of 25 $\mu g/m^3$ suggested by the WHO and the American Lung Association. However, the difference is marginal and as a result will unlikely lead to an increase in short-term health risks.

Lastly, using short-term mortality estimates from recent publications, hypothetical game day levels of $PM_{2.5}$ required to increase the risk of cardiovascular mortality, respiratory mortality and asthma symptoms from 0% to 1% are calculated assuming air quality standards are set to month and city-specific non-game day averages of $PM_{2.5}$.[30] These hypothetical game day levels are presented in Table 15. For October game days, on average $PM_{2.5}$ levels would have to be 105.5, 69 or 45.2 $\mu g/m^3$ in order to increase cardiovascular mortality, respiratory mortality or asthma symptoms, from

[30] See the notes in Table 15 for more details and for the cited literature.

a 0% to a 1% risk, respectively, with no notable deviations from these averages in terms of the by-city results. For November game days, the average $PM_{2.5}$ levels are similar with required game day levels of 110.4, 69.9 or 46.1 µg/m^3 in order to increase cardiovascular mortality, respiratory mortality or asthma symptoms, from a 0% to a 1% risk. These hypothetical game day pollution levels are on average 8, 5 and 3 times larger, respectively, when compared to observable maximum game day levels in October and November. Thus, cities experience a marginal change in pollution during game days in October and November but these changes are not expected to significantly or even marginally increase mortality risks.

2.8 Conclusion

This paper added to the broad literature of exploring how sporting events impact cities by investigating an additional, previously unexamined, externality—the environmental impact of hosting college football games. Four questions were considered: do cities experience a day-of effect; do cities also experience a commuting effect; do these effects change over the course of the regular season; and, do these effects pose serious health-risks? Results indicated that cities do experience both day-of and commuting effects. These effects tended to be negligible in September, marginally increasing in October and were largest in November. While the increase in pollution on and around game days were largest in November, predicted levels were below hazardous levels and hypothetical levels required to increase various mortality rates from 0% to 1% far exceeded observed maximum game day levels. This provided convincing evidence that the estimated marginal increase in daily pollution as a result of hosting a college football game was not hazardous and was not expected to increase mortality risks.

In an effort to continue to maintain below hazardous levels of game day pollution, cities should extend their traffic-management efforts to before, during and after game days. These efforts should intensify as the regular season progresses and particularly for themed games like “Homecoming” in October and “Rivalry Week” in November. Additionally, cities could broadcast public service announcements regarding the possibility of sudden air quality changes—albeit these changes may be small in absolute magnitude and predicted levels should be below hazardous levels. To contain cost, these announcements could be sent via local media outlets with a focus towards particularly susceptible cohorts including children, older adults and those with preexisting respiratory conditions such as asthma.

2.9 Tables

Table 4. Full Sample and by City Demographics.

	(1)			(2)					(3)
	Macroeconomic			**Employment**					**Commute**
	Pop.	*Inc.*	*U.R.*	*Manage*	*Serv.*	*Sales*	*Constr.*	*Trans.*	*Drive*
Full Sample:	24.4	2.2	8.6	37.0	19.9	25.4	7.3	10.4	18.6
By City:									
Akron, OH	19.9	2.0	12.9	27.7	21.6	26.1	7.8	16.7	21.0
Albuquerque, NM	54.6	2.6	6.3	39.0	18.0	25.4	9.1	8.4	21.2
Austin, TX	79.0	3.0	6.5	43.7	17.0	23.0	9.9	6.4	22.7
Baton Rouge, LA	22.9	2.3	8.8	34.9	20.6	26.4	7.6	10.5	20.7
Birmingham, AL	21.2	2.0	12.9	29.7	21.0	26.2	8.3	14.9	22.0
Bloomington, IN	8.0	1.8	7.5	46.8	21.9	20.1	3.4	7.9	15.6
Colorado Springs, CO	41.6	2.8	7.6	39.5	17.8	25.2	8.9	8.6	20.8
Columbia, SC	12.9	2.4	10.6	41.8	20.9	24.7	5.0	7.6	17.6
Columbus, OH	78.7	2.3	8.9	36.8	17.8	28.1	5.9	11.4	20.8
Durham, NC	22.8	2.7	8.4	47.9	16.5	20.7	7.8	7.0	21.3
Eugene, OR	15.6	2.5	8.2	40.7	17.5	26.3	6.3	9.2	16.8
Fresno, CA	49.5	2.0	12.4	29.0	20.6	27.2	11.3	11.9	20.8
Gainesville, FL	12.4	1.9	7.7	43.5	23.0	22.7	4.3	6.4	16.7
Hattiesburg, MS	4.6	1.8	11.2	32.5	22.2	24.2	8.3	12.9	17.7
Iowa City, IA	6.8	2.5	5.1	45.2	20.6	23.4	3.6	7.3	16.1
Kalamazoo, MI	7.4	1.8	15.6	32.0	24.7	26.1	5.5	11.7	17.9
Knoxville, TN	17.9	2.2	7.1	34.6	20.1	28.0	7.5	9.9	18.6
Lafayette, LA	12.1	2.8	5.9	37.4	18.3	27.0	9.1	8.2	20.4
Las Cruces, NM	9.8	2.0	8.5	37.5	21.5	24.6	9.3	7.1	16.7
Lincoln, NE	25.8	2.5	5.6	38.3	17.1	25.3	8.0	11.4	17.3

Logan, UT	4.8	1.8	6.4	33.4	18.5	26.5	6.8	14.8		13.6
Louisville, KY	59.7	2.5	9.6	33.0	17.1	26.2	7.8	15.9		21.8
Lubbock, TX	23.0	2.3	6.2	33.2	19.5	28.4	9.0	9.9		15.5
Madison, WI	23.3	3.0	5.6	49.5	16.3	23.0	3.8	7.4		18.7
Monroe, LA	4.9	2.0	10.8	30.3	23.6	29.6	6.1	10.4		17.6
Muncie, IN	7.0	1.7	15.2	27.6	25.5	26.6	5.9	14.4		17.2
Provo, UT	11.2	1.6	6.5	36.7	19.1	28.4	6.9	8.9		16.6
Salt Lake City, UT	18.6	2.6	6.5	40.6	16.8	24.0	7.6	11.0		19.2
State College, PA	4.2	1.3	9.3	45.5	22.6	23.0	3.9	5.1		14.6
Tucson, AZ	52.0	2.0	8.6	32.1	22.2	26.1	10.6	9.0		21.5
Tulsa, OK	39.2	2.6	7.0	33.8	17.5	26.4	10.2	12.0		17.9
Tuscaloosa, AL	9.0	2.1	7.2	31.9	20.0	24.6	8.3	15.2		16.3
Winston-Salem, NC	23.0	2.4	8.5	36.5	18.2	24.5	8.5	12.4		19.3

Source: U.S. Census Bureau 2006-2010 American Community Survey.
Notes: City population (Pop.) is measured in 10,000's. Real per capita income (Inc.) is measured in $10,000's. Unemployment rates (U.R.) are expressed as a percent. All employment variables are expressed as a percent. Lastly, average time to work (Drive) is measured in minutes.

Table 5. Hosted FBS Teams and Team Quality Measures.

	(1)	(2)	(3)	(4)
	FBS Team	*SW%*	*CW%*	*Att.*
Full Sample:		0.54	0.51	5.46
By City:				
Akron, OH	Akron Zips	0.08	0.12	1.02
Albuquerque, NM	New Mexico Lobos	0.08	0.12	2.09
Austin, TX	Texas Longhorns*	0.42	0.25	10.1
Baton Rouge, LA	Louisiana State Tigers*	0.85	0.75	9.70
Birmingham, AL	Alabama-Birmingham Blazers	0.33	0.38	2.07
Bloomington, IN	Indiana Hoosiers	0.42	0.12	4.22
Colorado Springs, CO	Air Force Falcons	0.69	0.62	4.05
Columbia, SC	South Carolina Gamecocks*	0.64	0.56	7.67
Columbus, OH	Ohio State Buckeyes*	0.92	0.88	10.6
Durham, NC	Duke Blue Devils	0.25	0.12	2.14
Eugene, OR	Oregon Ducks	0.92	1.00	5.94
Fresno, CA	Fresno State Bulldogs	0.62	0.62	3.41
Gainesville, FL	Florida Gators*	0.62	0.50	9.07
Hattiesburg, MS	S. Miss. Golden Eagles	0.62	0.62	2.83
Iowa City, IA	Iowa Hawkeyes	0.62	0.50	7.06
Kalamazoo, MI	Western Michigan Broncos	0.50	0.62	1.67
Knoxville, TN	Tennessee Volunteers*	0.46	0.38	9.98
Lafayette, LA	LA-Lafayette Ragin' Cajuns	0.25	0.38	1.74
Las Cruces, NM	New Mexico State Aggies	0.17	0.12	1.63
Lincoln, NE	Nebraska Cornhuskers*	0.71	0.67	8.61
Logan, UT	Utah State Aggies	0.33	0.25	1.79
Louisville, KY	Louisville Cardinals	0.54	0.43	5.06

Lubbock, TX	Texas Tech Red Raiders	0.62	0.38	5.75
Madison, WI	Wisconsin Badgers*	0.85	0.88	7.99
Monroe, LA	LA-Monroe Warhawks	0.42	0.50	5.57
Muncie, IN	Ball State Cardinals	0.33	0.38	0.99
Provo, UT	Brigham Young Cougars	0.54	0.62	6.20
Salt Lake City, UT	Utah Utes	0.77	0.88	4.55
State College, PA	Penn State Nittany Lions*	0.54	0.50	10.4
Tucson, AZ	Arizona Wildcats	0.54	0.44	5.53
Tulsa, OK	Tulsa Golden Hurricane	0.77	0.75	2.04
Tuscaloosa, AL	Alabama Crimson Tide*	0.77	0.62	10.2
Winston-Salen, NC	Wake Forest Demon Deacons	0.25	0.12	3.05

Source: NCAA.org.

Notes: SW% denotes a team's 2010-2011 season winning percentage expressed as a decimal. CW% denotes a team's 2010 conference winning percentage expressed as a decimal. Average attendance (Att.) is measured in 10,000's. A star, *, denotes teams that play in one of the top 20 largest stadiums by capacity in college football, see text for further details.

Table 6. Information on Pollution and Weather Monitors.

	(1)	(2)	(3)
	Numb. P.M.	*Distance: P.M. to Stadium*	*Distance: W.M. to Stadium*
Full Sample:	1.5	3.1	4.82
By City:			
Akron, OH	2	2.14	3.33
Albuquerque, NM	2	4.09	1.86
Austin, TX	1	1.85	7.65
Baton Rouge, LA	1	3.48	8.95
Birmingham, AL	2	4.05	6.84
Bloomington, IN	1	1.76	0.47
Colorado Springs, CO	1	10.3	15.4
Columbia, SC	1	1.31	0.7
Columbus, OH	3	4.71	7.4
Durham, NC	1	3.3	12.6
Eugene, OR	2	2.64	8.96
Fresno, CA	2	3.11	3.21
Gainesville, FL	2	4.36	5.27
Hattiesburg, MS	1	2.59	5.16
Iowa City, IA	1	2.46	4.16
Kalamazoo, MI	1	3.07	4.34
Knoxville, TN	4	4.17	5.37
Lafayette, LA	2	1.31	1.37
Las Cruces, NM	1	3.34	1.14
Lincoln, NE	1	1.31	3.03
Logan, UT	1	1.97	0.62
Louisville, KY	3	3.38	2.02
Lubbock, TX	1	1.35	5.92
Madison, WI	2	2.54	5.96
Monroe, LA	1	1.87	1.83

Muncie, IN	1	1.8	1.73
Provo, UT	1	0.53	12.6
Salt Lake City, UT	2	3.59	2.62
State College, PA	1	1.11	1.43
Tucson, AZ	2	6.62	6.75
Tulsa, OK	1	4.29	4.71
Tuscaloosa, AL	1	4.03	3.81
Winston-Salem, NC	2	3.85	1.81

Sources: Pollution data taken from the Air Quality System database; weather data taken from the Global Historical Climatology Network; football stadium locations were taken from Google Maps.
Note: Distance was measured in miles and calculated using the great-circle distance formula.

Table 7. Maximum Pollution on Game Days vs. Non-Game Days.

	A. Game Day		**B. Non-Game Day**	
	(1)	**(2)**	**(3)**	**(4)**
	N	*Max* $PM_{2.5}$	*N*	*Max* $PM_{2.5}$
Full Sample:	166	66	1936	47.1
By City:				
Akron, OH	6	14.5	85	31.6
Albuquerque, NM	6	6.6	83	13.7
Austin, TX	2	14.4	12	16.3
Baton Rouge, LA	9	23.9	76	22.7
Birmingham, AL	13	24.8	78	31.2*
Bloomington, IN	4	13	74	21.5
Colorado Springs, CO	3	7.7	27	9
Columbia, SC	7	17.4	82	22.7
Columbus, OH	5	22.2	25	21.1*
Durham, NC	4	16.3	25	20.5
Eugene, OR	6	11	84	19.4*
Fresno, CA	6	40.2	83	47.1*
Gainesville, FL	2	7.25	28	14.6
Hattiesburg, MS	3	13.7	24	20.5
Iowa City, IA	6	15.6	81	20.6*
Kalamazoo, MI	3	19.9	26	20.0*
Knoxville, TN	7	19.1	84	24*
Lafayette, LA	5	25.1	84	37.5
Las Cruces, NM	3	8	25	9
Lincoln, NE	3	8.7	27	11.9
Logan, UT	6	12.3	83	37.7
Louisville, KY	7	66	84	46*
Lubbock, TX	1	4.4	9	19.4
Madison, WI	7	16.6	84	27.1*
Monroe, LA	1	8.3	28	25.8
Muncie, IN	2	12	21	27.8
Provo, UT	4	8.2	76	13.4*
Salt Lake City, UT	6	30.1	77	20.6*
State College, PA	7	8.89	84	26.6*
Tucson, AZ	3	6.85	27	6.25
Tulsa, OK	6	15.4	83	23.3
Tuscaloosa, AL	7	18.3	84	27.2
Winston-Salem, NC	6	19	83	21

Note: A star, *, indicates that the maximum $PM_{2.5}$ level occurred within three days of a home football game.

Table 8. Differences in Pollution on Game Days vs. Non-Game Days by City.

	(1)	(2)	(3)	(4)	(5)
	Diff.	*SE*	*L.T.*	*T.T.*	*R.T.*
Full Sample:	0.79	0.49	0.94	0.11	0.06
By City:					
Akron, OH	-2.64	1.70	0.08	0.16	0.92
Albuquerque, NM	-0.35	0.57	0.28	0.56	0.72
Austin, TX	2.58	2.60	0.77	0.46	0.23
Baton Rouge, LA	1.76	2.09	0.79	0.42	0.21
Birmingham, AL	-0.07	1.69	0.48	0.97	0.52
Bloomington, IN	-0.96	2.00	0.33	0.66	0.67
Colorado Springs, CO	0.95	0.68	0.88	0.23	0.12
Columbia, SC	2.86	1.61	0.94	0.12	0.06
Columbus, OH	2.02	3.05	0.73	0.54	0.27
Durham, NC	1.47	2.09	0.74	0.52	0.26
Eugene, OR	-1.38	1.44	0.19	0.37	0.81
Fresno, CA	6.06	5.50	0.84	0.32	0.16
Gainesville, FL	-1.43	0.86	0.07	0.14	0.93
Hattiesburg, MS	-1.78	1.95	0.21	0.42	0.79
Iowa City, IA	-0.96	2.06	0.33	0.66	0.67
Kalamazoo, MI	1.21	5.39	0.58	0.84	0.42
Knoxville, TN	-0.51	1.76	0.39	0.78	0.61
Lafayette, LA	1.95	3.64	0.69	0.62	0.31
Las Cruces, NM	0.69	1.58	0.65	0.7	0.35
Lincoln, NE	0.58	1.53	0.63	0.74	0.37
Logan, UT	0.71	1.74	0.65	0.7	0.35
Louisville, KY	7.78	8.10	0.81	0.37	0.19
Lubbock, TX	N.A.				
Madison, WI	-0.20	1.70	0.45	0.91	0.55
Monroe, LA	N.A.				
Muncie, IN	1.23	1.69	0.76	0.48	0.24
Provo, UT	-0.42	1.08	0.36	0.72	0.64
Salt Lake City, UT	3.26	4.11	0.77	0.46	0.23
State College, PA	-2.29	0.98	0.02	0.04	0.98
Tucson, AZ	1.11	0.65	0.89	0.21	0.11
Tulsa, OK	0.31	1.41	0.58	0.83	0.42
Tuscaloosa, AL	-3.03	1.56	0.04	0.09	0.96
Winston-Salem, NC	1.20	2.35	0.69	0.63	0.31

Notes: Diff. denotes the difference in average $PM_{2.5}$ levels on game days versus non-game days for the full sample and by city. SE denotes the corresponding standard error. L.T., T.T. and R.T. provide the p-values for whether the reported difference in means is less than, differs from or exceeds zero, respectively.

Table 9. Differences in Pollution on Game Days vs. Non-Game Days by Month and City Population.

	(1)	(2)	(3)	(4)	(5)
	Diff.	*SE*	*L.T.*	*T.T.*	*R.T.*
By Month:					
September	-0.77	0.68	0.13	0.26	0.87
October	1.69	0.83	0.98	0.04	0.02
November	2.11	1.06	0.98	0.05	0.02
By Population:					
Small Cities	-0.78	0.69	0.13	0.26	0.87
Large Cities	1.24	0.77	0.95	0.11	0.05
By Population & Month:					
Small Cities:					
September	-1.02	0.92	0.13	0.27	0.87
October	0.08	1.14	0.53	0.94	0.47
November	1.39	1.32	0.85	0.3	0.15
Large Cities:					
September	-0.52	1.02	0.31	0.61	0.69
October	2.97	1.22	0.99	0.02	0.01
November	2.48	1.65	0.93	0.13	0.07

Notes: Diff. denotes the difference in average $PM_{2.5}$ levels on game days versus non-game days by month and by city population. SE denotes the corresponding standard error. L.T., T.T. and R.T. provide the p-values for whether the reported difference in means is less than, differs from or exceeds zero, respectively. Small cities are defined as cities with population levels less than or equal to the sample median of 186,440; large cities are similarly defined but have populations exceeding the sample median.

Table 10. Average Attendance and Weather by Month.

	(1)	(2)	(3)	(4)
	Attendance	*Away Games*	*Temperature*	*Precipitation*
Full Sample:	50.9	0.43	54.8	2.23
	(33.9)	(0.50)	(16.8)	(8.10)
	{167}	{292}	{2660}	{2776}
By Month:				
September	52.9	0.31	72.1	1.70
	(33.7)	(0.46)	(8.55)	(6.14)
	{66}	{95}	{669}	{698}
October	48.0	0.47	60.9	1.94
	(33.6)	(0.50)	(8.88)	(7.49)
	{52}	{99}	{684}	{715}
November	51.8	0.44	48.7	3.03
	(35.3)	(0.50)	(11.2)	(9.06)
	{48}	{85}	{663}	{687}

Notes: Average attendance (Attendance) is measured in 1,000's. Away Games measures the proportion of times teams played an away game or a game played at a neutral site. Temperature is the midrange of minimum and maximum daily temperature readings and is measured in Fahrenheit. Precipitation is measured in millimeters. Legend for all cells: average / (standard deviation) / {number of observations}.

Table 11. Fixed-Effects Estimates of the Day-of Effect of Hosting a Game.

	(1)	(2)	(3)	(4)
	Fixed-Effects	*Add Weather*	*Add Time Trend*	*Add City-Specific Time Trends*
Host Game	0.72	0.44	0.46	0.45
	(0.51)	(0.50)	(0.51)	(0.51)
Temperature		1.30**	1.42**	1.65***
		(0.53)	(0.54)	(0.55)
Precipitation		-0.14***	-0.14***	-0.15***
		(0.02)	(0.02)	(0.02)
Linear Time Trend			0.04	
			(0.02)	
R-squared	0.24	0.28	0.28	0.36
Observations	2102	2016	2016	2016

Notes: White-robust standard errors clustered by city are shown in parentheses. For all columns, the dependent variable is average $PM_{2.5}$. Levels of statistical significance are indicated as follows: *10%; **5%; and, ***1%. For Temperature, a one unit change is 10 degrees Fahrenheit and Precipitation is measured in millimeters.

Table 12. Fixed-Effects Estimates of the Commuting and Day-of Effects of Hosting a Game.

	(1)	(2)	(3)	(4)
	Fixed-Effects	*Add Weather*	*Add Time Trend*	*Add City-Specific Time Trends*
Day Before	0.85*	0.66	0.75	0.72
	(0.47)	(0.44)	(0.46)	(0.44)
Host Game	0.82	0.52	0.55	0.52
	(0.57)	(0.55)	(0.56)	(0.56)
Day After	0.26	0.19	0.2	0.12
	(0.42)	(0.42)	(0.42)	(0.39)
Temperature		1.30**	1.42**	1.65***
		(0.53)	(0.54)	(0.55)
Precipitation		-0.14***	-0.14***	-0.14***
		(0.02)	(0.02)	(0.02)
Linear Time Trend			0.04	
			(0.03)	
Observations	2102	2016	2016	2016
R-squared	0.24	0.28	0.29	0.36

Notes: White-robust standard errors clustered by city are shown in parentheses. For all columns, the dependent variable is average $PM_{2.5}$. Levels of statistical significance are indicated as follows: *10%; **5%; and, ***1%. For Temperature, a one unit change is 10 degrees Fahrenheit and Precipitation is measured in millimeters.

Table 13. Fixed-Effects Estimates for Month-Specific Commuting and Day-of Effects for Hosting a Game.

	(1)	**(2)**	**(3)**	**(4)**
	Fixed-Effects	*Add Weather*	*Add Time Trend*	*Add City-Specific Time Trends*
Day Before	0.19	-0.41	0.28	-0.11
	(0.60)	(0.66)	(0.65)	(0.55)
Host Game	-0.86	-1.58**	-0.95	-1.14
	(0.61)	(0.64)	(0.67)	(0.69)
Day After	-0.91	-1.42**	-0.63	-0.98**
	(0.59)	(0.64)	(0.59)	(0.48)
Day Before*Oct.	0.82	0.76	0.22	0.58
	(1.30)	(1.32)	(1.35)	(1.24)
Host Game*Oct.	2.45	2.62*	2.04	2.28
	(1.52)	(1.49)	(1.47)	(1.44)
Day After*Oct.	1.91	1.98*	1.22	1.58
	(1.17)	(1.15)	(1.21)	(1.08)
Day Before*Nov.	1.21	2.54**	1.42	2.10*
	(1.08)	(1.17)	(1.05)	(1.22)
Host Game*Nov.	3.02***	4.09***	2.89**	3.26**
	(1.05)	(1.15)	(1.07)	(1.20)
Day After*Nov.	1.80**	2.83**	1.35	1.92**
	(0.76)	(1.08)	(0.82)	(0.86)
Temperature		0.31	1.23**	1.58***
		(0.30)	(0.46)	(0.52)
Precipitation		-0.12***	-0.14***	-0.14***
		(0.01)	(0.02)	(0.02)
Linear Time Trend			0.06**	
			(0.02)	
Observations	2102	2016	2016	2016
R-squared	0.24	0.26	0.28	0.37

Notes: White-robust standard errors clustered by city are shown in parentheses. For all columns, the dependent variable is average $PM_{2.5}$. Levels of statistical significance are indicated as follows: *10%; **5%; and, ***1%. For Temperature, a one unit change is 10 degrees Fahrenheit and Precipitation is measured in millimeters. The base month is September.

Table 14. Percent Change and Predicted Levels in Pollution for Game Days in October and November.

	A. Game in October		B. Game in November	
	(1)	(2)	(3)	(4)
	Percent Change	*Predicted Level*	*Percent Change*	*Predicted Level*
Full Sample:	24.0	11.79	31.3	13.66
By City:				
Akron, OH	19.7	13.88	22.0	18.06
Albuquerque, NM	42.1	7.7	70.9	7.86
Austin, TX	19.5	13.98	48.4	10.00
Baton Rouge, LA	18.8	14.38	36.5	12.19
Birmingham, AL	16.1	16.48	27.2	15.26
Bloomington, IN	22.9	12.25	26.5	15.56
Colorado Springs, CO	42.9	7.59	66.8	8.14
Columbia, SC	19.7	13.88	33.9	12.87
Columbus, OH	23.4	12.01	26.5	15.56
Durham, NC	29.4	10.04	35.8	12.37
Eugene, OR	24	11.79	46.2	10.32
Fresno, CA	21.5	12.88	13.8	26.96
Gainesville, FL	22.8	12.28	51.5	9.59
Hattiesburg, MS	18.4	14.68	29.1	14.46
Iowa City, IA	29.1	10.11	33.9	12.87
Kalamazoo, MI	31.7	9.47	25.9	15.86
Knoxville, TN	21.5	12.88	27.4	15.16
Lafayette, LA	15.0	17.48	28.1	14.86
Las Cruces, NM	42.5	7.64	59.4	8.75
Lincoln, NE	37.3	8.39	55.6	9.12
Logan, UT	36.2	8.57	36.2	12.27
Louisville, KY	19.7	13.88	26.3	15.66
Lubbock, TX	41.0	7.84	N.A.	N.A.
Madison, WI	27.6	10.53	25.9	15.86
Monroe, LA	15.6	16.88	44.5	10.58
Muncie, IN	29.0	10.15	23.1	17.36
Provo, UT	39.0	8.12	47.7	10.09
Salt Lake City, UT	36.4	8.55	41.6	11.1
State College, PA	34.7	8.86	33.5	12.99

Tucson, AZ	51.5	6.71	70.0	7.92
Tulsa, OK	26.4	10.92	44.1	10.66
Tuscaloosa, AL	16.2	16.38	23.5	17.16
Winston-Salem, NC	27.2	10.65	34.8	12.63

*Not*es: Percent change is the ratio of the estimated month-specific day-of effect in Table 2.10, column (4) and the month-specific average $PM_{2.5}$ level on non-game days for the sample and by city. Similarly, the predicted level is the estimated month-specific day-of effect in Table 2.10, column (4) plus the month-specific average $PM_{2.5}$ level on non-game days for the sample and by city.

Table 15. Levels of Pollution Required to Increase Mortality Risks.

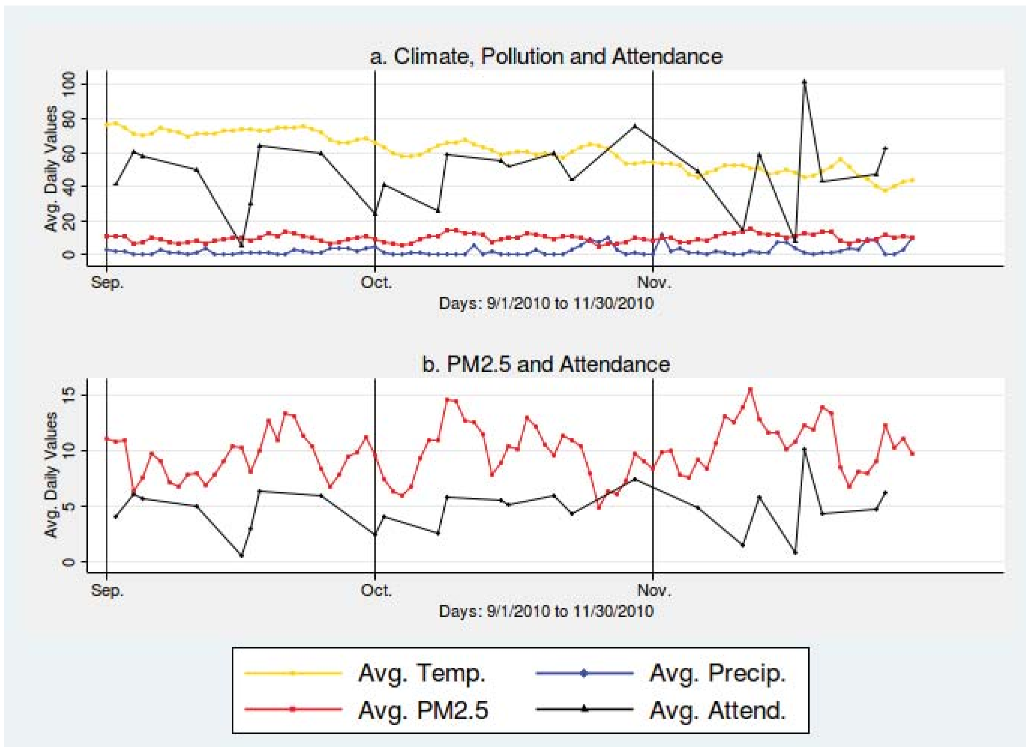
	A. Game in October			**B. Game in November**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Cardio.	*Resp.*	*Asthma*	*Cardio.*	*Resp.*	*Asthma*
Full Sample:	109.5	69.0	45.2	110.4	69.9	46.1
By City:						
Akron, OH	111.6	71.1	47.3	114.8	74.3	50.5
Albuquerque, NM	105.4	64.9	41.1	104.6	64.1	40.3
Austin, TX	111.7	71.2	47.4	106.7	66.3	42.5
Baton Rouge, LA	112.1	71.6	47.8	108.9	68.5	44.6
Birmingham, AL	114.2	73.7	49.9	112.0	71.5	47.7
Bloomington, IN	110.0	69.5	45.7	112.3	71.8	48.0
Colorado Springs, CO	105.3	64.8	41.0	104.9	64.4	40.6
Columbia, SC	111.6	71.1	47.3	109.6	69.1	45.3
Columbus, OH	109.7	69.3	45.4	112.3	71.8	48.0
Durham, NC	107.8	67.3	43.5	109.1	68.6	44.8
Eugene, OR	109.5	69	45.2	107.1	66.6	42.8
Fresno, CA	110.6	70.1	46.3	123.7	83.2	59.4
Gainesville, FL	110.0	69.5	45.7	106.3	65.9	42.0
Hattiesburg, MS	112.4	71.9	48.1	111.2	70.7	46.9
Iowa City, IA	107.8	67.4	43.5	109.6	69.1	45.3
Kalamazoo, MI	107.2	66.7	42.9	112.6	72.1	48.3
Knoxville, TN	110.6	70.1	46.3	111.9	71.4	47.6
Lafayette, LA	115.2	74.7	50.9	111.6	71.1	47.3
Las Cruces, NM	105.4	64.9	41.1	105.5	65.0	41.2
Lincoln, NE	106.1	65.6	41.8	105.9	65.4	41.6
Logan, UT	106.3	65.8	42.0	109	68.5	44.7
Louisville, KY	111.6	71.1	47.3	112.4	71.9	48.1
Lubbock, TX	105.6	65.1	41.3	N.A.	N.A.	N.A.

Madison, WI	108.3	67.8	44.0	112.6	72.1	48.3
Monroe, LA	114.6	74.1	50.3	107.3	66.8	43.0
Muncie, IN	107.9	67.4	43.6	114.1	73.6	49.8
Provo, UT	105.8	65.4	41.6	106.8	66.4	42.5
Salt Lake City, UT	106.3	65.8	42.0	107.8	67.4	43.6
State College, PA	106.6	66.1	42.3	109.7	69.3	45.4
Tucson, AZ	104.4	64	40.1	104.7	64.2	40.4
Tulsa, OK	108.6	68.2	44.4	107.4	66.9	43.1
Tuscaloosa, AL	114.1	73.6	49.8	113.9	73.4	49.6
Winston-Salem, NC	108.4	67.9	44.1	109.4	68.9	45.1
Average Factor:	8.8	5.5	3.6	8.4	5.2	3.4

Notes: Above are hypothetical $PM_{2.5}$ levels required to increase the risk of cardiovascular mortality, respiratory mortality and asthma from 0% to 1% assuming month and city-specific non-game day averages of $PM_{2.5}$ established air quality standards. The calculation follows the same approach in WHO (2005), specifically Table 2.2. See Cohen et al. (2004) for more details. Point-estimates for short-term mortality effects were set at 1% for cardiovascular mortality (Brook et al. 2010), 1.68% for respiratory mortality (Anderson et al. 2012) and 2.8% for asthma symptoms in children (Weinmayr et al. 2010). Lastly, Average Factor denotes the average magnitude of these hypothetical levels relative to the city-specific sample maximum of $PM_{2.5}$ on game days.

2.10 Figures

Figure 5. Day-Level Averages over Time.



CHAPTER III

ESSAY 3: FEWER CALORIES, MORE DRINKERS: DID THE INITIAL ULTRA-LIGHT BEERS ATTRACT NEW DRINKERS?

3.1 Abstract

This paper estimates the demand for beer in the U.S. from 2001 until 2006 using a multinomial logit discrete-choice model of product differentiation.[1] Using grocery store scanner data from Information Resources Incorporated (IRI), demand estimates are used to evaluate the extent to which two new beer brands—Michelob Ultra and Bud Select—attracted new drinkers. This unexplored aspect to new brands has various public health implications regarding the over-consumption of alcohol. Counterfactual results based on a single market suggest that 68% of sales of Bud Select and 74% of sales of Michelob Ultra were due to new drinkers. Additionally, new drinkers of Bud Select seemed to prefer larger package sizes—specifically 12-packs over 6-packs—whereas the reverse held for Michelob Ultra. Lastly, a number of suggestions for future work are provided.

3.2 Introduction

Recent studies examining the private welfare effects of new goods have typically examined innocuous products—in the sense that they fail to impose any additional externalities once consumed. Most motivate the study of new brands with policy

[1] All estimates and analyses in this paper based on SymphonyIRI Group, Inc. data are by the author and not by SymphonyIRI Group, Inc.

implications towards properly calculating price indexes. Recent examples include examining the welfare impacts of new brands of ready-to-eat cereal (Hausman 1997, Nevo 2003), phone directories (Rysman 2004, Ackerberg et al. 2005), online newspapers (Gentzkow 2007), tomography scanners (Trajtenberg 1989), cellular telephones (Hausman 1999) and minivans (Petrin 2002). However, the introduction of new beer brands is not innocuous given the various public health concerns regarding gate-way effects in teens (Kirby et al. 2012), risky-behavior (Pedrelli et al. 2011, Rehm et al. 2012 and Snipes et al. 2013), over-consumption leading to cirrhosis and various cancers (Temple 2012) and the various effects on crime (Carpenter et al. 2010, Lovenheim et al. 2011, Popovici et al. 2012, Hansen 2013).

This paper adds to the literature by examining the extent to which two new beer brands—Michelob Ultra and Bud Select—attracted new drinkers and if these new drinkers purchased larger package-sizes.[2] These subtle and unexplored aspects to new brands have possible public health implications regarding the consequences of over-consumption of alcohol and possible gate-way effects. To examine the effects of new beer brands in terms of attracting new drinkers, a multinomial logit discrete-choice demand model of product differentiation is estimated using grocery store scanner data from Information Resources Incorporated (IRI) from 2001 until 2006.[3] Thus,

[2]This latter aspect builds on the previous work of Bray et al. (2007, 2009) who analyzed the advertisements of larger package-sizes and degree of substitution towards larger package sizes within brands. The analysis here is more flexible—permitting substitution within brand, across brand and across package sizes. Additionally, recovering consumer tastes permits policy counterfactuals.

[3]The use of scanner data for analyzing the U.S. beer industry is strongly suggested by Ruhm et al. (2011) and is common in discrete-choice demand estimation. Notable previous U.S. beer product differentiated demand studies include Hausman et al. (1994), Rojas et al. (2008) and Rojas (2008) which examine hypothetical mergers, own and cross-advertising elasticities and pricing behavior of major beer companies, respectively. In addition, all make use, in some form or another, of the almost ideal demand system by Deaton and Muellbauer (1980a, 1980b)—typically used for its theoretical properties such as adding up restrictions, symmetry and homogeneity. Econometrically, such properties reduce the number of demand parameters to be estimated thereby overcoming the classic

consumers take into account all observed and unobserved product characteristics to optimally choose a beer brand and package size to maximize utility given the available choice-set of beer brands in a given market and the option of an outside good. Demand estimates along with an assumption regarding firm behavior are then used to consider counterfactuals.[4]

In the context of the demand model, the intuition for the counterfactuals is to measure the number of consumers in a given market that choose Michelob Ultra (and similarly for Bud Select) when it is available in the choice-set. Additionally, the package size choice can also be determined. After identifying this set of consumers, counterfactuals involve removing these brands from the choice set and tracking how many of these former purchasers of Michelob Ultra (and similarly for Bud Select) switched to the outside good—all other goods except beer. Given the data consists of grocery store beer purchases, the outside good serves as the non-alcohol option. Hence, analyzing how many former purchasers switch to the outside good once the brand is removed acts as the measure of how strong a new brand attracted new drinkers.

dimensionality problem in demand estimation. However, Nevo (2011) notes that the derivation of such demand models usually assumes that consumers purchase a positive amount of all products. This is problematic in product differentiated demand studies since demand is analyzed across the various brands for a single product. In the case of beer, it may be reasonable to assume consumers buy more than one brand of beer but it would be unreasonable to assume consumers buy some amount of all brands of beer. Lastly, Hellerstein (2008) and Goldberg et al. (2010) estimate own and cross-price via a random-coefficients logit demand model for the U.S. beer market with applications in international trade—specifically how exchange rate shocks impact consumer and firm behavior; both used scanner data from Dominick's Finer Foods.

[4]The convention in the literature has been to model supply assuming firms behave as Bertrand-Nash competitors; thus, multi-product firms maximize profits by competing in prices. The counterfactuals considered mimic that of Petrin (2002) and, while the convention has been to model firm behavior as Bertrand-Nash competitors, Rojas (2008) does find empirical evidence in support of this pricing behavior specifically for the U.S. beer industry.

Counterfactual results based on a simulation involving 50,000 hypothetical consumers drawn from a type 1 extreme value distribution suggest sales were overwhelmingly generated by new drinkers. For a single market in 2003, the estimated number of consumers who chose Michelob Ultra when it was in the choice-set was 222 with 156 choosing 6-packs and 66 choosing 12-packs. When Michelob Ultra was excluded, 165 consumers switched to the outside good. Hence, for this single market, 74% of Michelob Ultra sales were generated by new drinkers and 26% were generated by existing drinkers.

For the same single market but in 2005, the estimated number of consumers who chose Bud Select when available in the choice-set was 69. These consumers favored larger package sizes with 60 choosing 12-packs and 9 choosing 6-packs. When Bud Select was excluded, 47 consumers switched to the outside good. Hence, for this single market, 68% of Bud Select sales were generated by new drinkers and 32% were generated by existing drinkers.

Given the various public health concerns regarding alcohol consumption, these results appear alarming. However, there are many notable limitations. First, the logit demand model does impose substantial restrictions on the own- and cross-price effects, firm mark-ups and mean utilities. Second, the counterfactuals considered were only for one market for each new brand and ignored any changes in brand-ownership as a result of mergers for simplicity. Third, a variety of alternative approaches and sensitivity analysis should be considered in future work.[5]

The remainder of this paper is organized as follows. Section 3.3 reviews the U.S. beer industry with a focus on light and ultra-light beers. Section 3.4 analyzes

[5] The details of which are discussed in the conclusion.

household-level data to describe company loyalty—a proxy for brand loyalty, package-size preferences and shopping versus drinking rates. The demand and supply model is discussed in Section 3.5. The data used in estimation is discussed in Section 3.6. Estimation and counterfactuals are discussed in Section 3.7. Lastly, a summary of key findings, important limitations and areas for future work conclude the paper.

3.3 U.S. Beer Industry and Ultra-light Beers

The U.S. beer industry is highly concentrated among three firms A-B, Miller, and Coors which account for 50.4%, 18.6% and 10.8% of the U.S. market based on volume, respectively.[6] Each of these three firms produces a large number of brands. For example, A-B produces over 200 brands whereas Coors produces over 40 brands. In fact, brand proliferation has been historically active in the light-beer category and more recently in the ultra-light beer category.

The light beer innovator was Miller with their introduction of Miller Lite in 1974. In terms of sales, Miller Lite was a tremendous success selling 400 thousand barrels in 1974 and 6.8 million barrels in 1977. Competitors soon launched their light beer brands in response to the success of Miller Lite. Coors launched Coors Light in 1978 selling 500 thousand barrels and experienced rapid growth shortly thereafter selling 3.1 million barrels in 1981. A-B launched Bud Light in 1982 selling 185 thousand

[6] This section is in large part based on a variety of industry and other available publications. Specifically, information on market shares and number of brands produced by major suppliers is from Datamonitor USA (2009, 2010). The review of the introduction of light beers during the 1970s and 1980s and their sales is from Tremblay and Tremblay (2005). The success of Bud Light by 2001, the rankings of top brands and all remaining sales figures both for light beers and ultra light beers are from various issues of Adams Beer Handbook (2001-2007, 2011). Beer characteristics, such as calories, carbohydrates and alcohol content per serving were taken from www.beer100.com. Remaining information—like brand history and brand fact sheets—was taken from official company websites.

barrels and similarly experienced rapid growth by selling 4.25 million barrels in 1984.[7] The popularity of Bud Light continued, surpassing its immediate rival brands of Miller Lite and Coors Light in 1994. In fact, Bud Light surpassed A-B leading beer Budweiser in 2001 and has since been not only the leading brand in the light beer category but the number one selling brand overall. By 2001, market share of light beer reached approximately 45% and 6 of the top 10 selling brands were light beers.

With continued success in the light beer category, A-B initiated the next round of brand proliferation with the introduction of ultra-light beers which have fewer calories and fewer carbohydrates per serving relative to light beers. In 2002, A-B launched Michelob Ultra—an ultra-light beer featuring 95 calories and 2.6 grams of carbohydrates per 12 ounce serving. The advertising campaign, which emphasized an active lifestyle coupled with a beer that has lower carbs, was perhaps influenced by the growing popularity of the low-carb Atkins Diet. Alternatively, the advent of ultra-light beers could have also been an attempt to boost sales and market share by appealing more towards females and in general towards younger, more health conscious consumers.

Michelob Ultra showed strong initial success, selling 5.9 million cases in 2002 and 41.5 million cases in 2003. In 2004, Coors responding by introducing Aspen Edge an ultra-beer featuring 94 calories and 2.6 grams of carbohydrates per 12 ounce serving; however, it was discontinued by 2005. In 2005, A-B produced two additional ultra-light beers, Bud Select and Michelob Ultra Amber.[8] Bud Select sold 32.4 million cases

[7] It is important to note that during the 1970s A-B did introduce two light-category beers: Natural Light in 1977 and Michelob Light in 1978. Both brands, while successful, failed to follow the sales pattern of rapid initial growth like that of Miller Lite and Coors Light. For more on the exact number of units sold for both Natural Light and Michelob Light, see Chapter 6 of Tremblay and Tremblay (2005).

[8] The latter was the first in many Michelob Ultra varieties, including: Michelob Ultra Lime Cactus

in its first year but sales dropped slightly to 30 million cases in 2006. The innovator of light beers Miller was the last to respond with the introduction of MGD 64 in 2008 with the advertising slogan, "As light as it gets." As the name implies, MGD 64 has 64 calories with 2.4 grams of carbohydrates per 12 ounce serving. By 2009, A-B retaliated with Bud Select 55—an ultra-light beer with 55 calories and 1.9 grams of carbohydrates per 12 ounce serving with the advertising slogan, "The lightest beer in the world."

The initial wave of ultra-light beers was generally successful. By 2006, Michelob Ultra and Bud Select ranked 13$^{\text{th}}$ and 15$^{\text{th}}$ respectively for overall top selling brands. However, the fact that they were both introduced by A-B, which produces many other light-beer brands, leads to the possibility of cannibalization. Sales data may provide evidence that Michelob Ultra cannibalized Michelob Light considerably.

Table 16 provides data on cases sold of select light and ultra-light beers from 2001 to 2006. Overall, the introduction of ultra-light beers by A-B failed to substantially alter Miller Lite and Coors Light sales. However, comparing sales of Michelob Light versus Michelob Ultra may provide evidence of brand cannibalization. In 2001, Michelob Light sold 40 million cases. Shortly after the introduction of Michelob Ultra, Michelob Light sales dropped slightly to 32 million in 2003. By 2006, Michelob Ultra sold 42 million cases; whereas for Michelob Light, sales further declined to 21 million cases.

---

introduced in 2007, Michelob Ultra Pomegranate Raspberry introduced in 2007 and Michelob Ultra Dragon Fruit Peach introduced in 2010. The introduction of these flavored beer varieties by A-B may have induced Miller to introduce Miller Chill in 2007, which A-B countered with Bud Light Lime in 2008. While it is interesting to consider the impacts of these flavored beers, the IRI data used in this analysis is from 2001 to 2006. Thus, due to these data limitations only the initial ultra-light beers can be examined.

Additionally, it is interesting to note that through 2006 the ultra-light beer varieties failed to cannibalize Bud Light sales. In fact, the A-B brands collectively sold significantly more cases relative to the competing brands of Miller Lite and Coors Light. The strong performance by A-B in 2006 may indicate that ultra-light beers attracted new customers to a variety of A-B products but failed to switch customers away from rival brand in favor of A-B brands. Thus while cannibalization may have occurred, it is possible that the ultra-light brands increased profits for A-B by attracting entirely new customers.

As an alternative explanation, brand cannibalization may not have occurred at all as the changes in light beer sales could be explained by specific changes in company expenditures towards advertising. Analyzing the same brands as Table 16, Table 17 provides information on advertising expenditure shares by company, the overall beer industry and the overall alcohol industry. Notice, advertising for Michelob Light declined considerably and persistently while advertising for Michelob Ultra remained relatively high but ultimately declined with the introduction of Bud Select. Thus, the decline in sales for Michelob Light may have in part been attributable to the decline in advertising. Similarly, the increase in sales for Michelob Ultra from 2002 to 2004 and decrease in sales from 2004 to 2005 may be due to a corresponding increase then decrease in advertising. Finally, steady sales for Bud Select from 2005 to 2006 may have been due to a strong commitment to advertising by A-B.

The continued growth of Bud Light from 2001 to 2006 in spite of new similar light beers launched by A-B was also coupled with strong advertising as it accounted for approximately one-third of A-B's advertising expenditure. More striking is the response of A-B's main competitors. In 2005, Miller and Coors had advertising shares of 70% and 87% for their respective major brands of Miller Lite and Coors Light. These levels could have been in response to A-B's success with Michelob Ultra and the initial launch of Bud Select. All of which would be consistent with the notion that A-B used increases in advertisement expenditures to mute brand cannibalization effects and competitors used increases in advertisement expenditures to protect their major brands from possible business stealing effects as A-B launched new light beer brands.

Also, the returns to advertisement expenditure by brand in the form of increased sales appear to favor A-B the most. Notice each had similar levels of total brand-level advertisement expenditure as each brand had on average an approximate 8% and 12% advertising expenditure share relative to total advertisement expenditure of the alcohol industry and beer industry, respectively. The dramatic difference occurs in total advertisement expenditure by company. On average from 1999 to 2006, A-B spent more on total advertising relative to Miller and Coors combined. Hence, the low by-brand shares for Bud Light in Table 17 relative to Miller Lite and Coors Light in spite of similar total brand-level advertising expenditure.

While the introduction of a new product is generally private welfare enhancing, the introduction of a new beer brand has various possible welfare effects given public health concerns regarding alcohol consumption. Typically a new brand will increase consumer surplus because of the additional value from having an additional product choice—either directly or indirectly via the impact of prices on existing products—or because product attributes are closer to a consumer's ideal brand or because of some combination of both effects. However, the specific characteristics of the initial ultra-light beers of Michelob Ultra and Bud Select can change private consumer surplus for much the same reasons but can also results in possible positive and negative health effects.

If Michelob Ultra and Bud Select cannibalized existing A-B brands or caused existing beer drinkers of rival brands to switch, there is a potential positive health effect as beer drinkers switch away from higher-calorie, higher-carbohydrate brands in favor of the relatively lower-calorie, lower-carbohydrate brands of Michelob Ultra or Bud Select. The reduction in calories and carbohydrates, ceteris paribus, would represent a marginal health benefit from a total calories consumed perspective. To judge the magnitude of this effect, Table 18 lists the basic product characteristics of popular light and ultra-light beers including calories, carbohydrates and alcohol content per 12 ounces. The initial ultra-light beers of Michelob Ultra and Bud Select offered a large reduction in carbohydrates, only a slight reduction in calories and an unchanged alcohol content by volume relative to popular light beer choices of Bud Light and Coors Light, but not Miller Lite. In fact, Miller Lite has comparable calories, carbohydrates and alcohol content as the initial ultra-light beers of Michelob Ultra and Bud Select.[9] Thus, the magnitude of a possible health benefit from a total calories perspective is likely to be very small.

If, on the other hand, Michelob Ultra and Bud Select attracted new consumers, there are various negative effects. Specifically, if these initial ultra-light beers attracted new drinkers–consumers who would not have purchased beer or alcohol in general in the absence of these new brands, then these consumers would consume more alcohol, ceteris paribus. If consumed in small amounts, the positive health effects of alcohol consumption include decreasing all-cause mortality, decreasing stroke

[9]Dramatic decreases in calories, carbohydrates and alcohol were associated with MGD 64 and Bud Select 55. However, due to data limitations, they are not further examined in this analysis.

risk, reducing the risk of coronary heart disease and heart attacks and decreasing the prevalence and incidence of diabetes.[10] However, if consumed in excess, the negative health effects of alcohol include cirrhosis and an increase risk of various cancers—colorectal, breast, mouth and pharynx, esophagus and liver.[11] Also, since the television ads for Michelob Ultra typically featured young active men and women and at the time a popular spokesman to all age groups, the targeted demographic of the brand seemed to be young, active and possibly health-conscience men and women. To the extent these brands attracted teenagers or young adult, negative public health concerns include a potential gateway effect where teenage drinking leads to further tobacco or drug (Kirby et al. 2012) and various high-risk behaviors in college students (Pedrelli et al. 2011 and Snipes et al. 2013).[12] Lastly, local public safety concerns include accidents, injuries and deaths caused by drunk drivers. In fact in 2006, alcohol was a factor in 40% of all fatal accidents in the U.S.[13]

3.4 Household Preferences

To complement the previous section, this section describes some key determinants of demand by analyzing household panel data from 2001 until 2006 for two markets: Eau Claire, WI and Pittsfield, MA.[14] Of particular interest is the degree of loyalty for products made by particular companies and the degree of loyalty for specific package

[10] See O'Keefe et al. (2007) and the additional references of this vast literature in the meta-analysis of 44 studies across 30 years by Roerecke et al. (2012a).

[11] See Temple (2012).

[12] See Rehm et al. (2012) for a meta-analysis of 12 studies and for additional references examining the relationship between alcohol consumption and risky-behavior with a focus on unprotected sex. For a recent study examining the impact alcohol advertising has on teen drinking, see Morgenstern et al. (2011).

[13] See Temple (2012).

[14] The household-level panel data was licensed from Information Resources Incorporated (IRI) where for each year 5,921 household beer purchases are recorded. For more information about the IRI data, see Bronnenberg et al. (2008).

sizes. The degree of loyalty should be informative regarding cannibalization versus business stealing effects. In particular, the cannibalization effect of a new brand would be relatively large when households are relatively more loyal to rival brands. In this case, the initial ultra-light beers would offer existing beer drinkers a small positive health benefit from a total-calories perspective. However, if preferences for rival brands are weak or if sales to new drinkers dominate any cannibalization effects because of effective advertising, then the launch of the initial ultra-light beers may increase the likelihood of various negative effects.

For package size preferences, available household-level data are used to investigate the possibility that households may purchase different package sizes but drink at similar rates. If true, this would imply larger package sizes facilitate storage which in turn has implications regarding demand estimation. If false and in particular if households purchase different package sizes but those who purchase larger package sizes drink at relatively higher rates, then this would provide evidence of immediate consumption which has implications for estimating demand. It would also have further health implications as any impacts on health from new brands attracting new drinkers would in turn depend on package size preferences. In particular, the new brands of the initial ultra-light beers may have been successful at attracting new beer drinkers but only for smaller package sizes. This could mute the negative health implications from the over-consumption of alcohol and possibly increase the positive health benefits associated with the consumption of alcohol in small amounts. If on the other hand the new brands attracted new drinkers who in turn have strong preferences for larger package sizes, then the negative health effects from the over-consumption of alcohol would likely dominate.

Average proportions of household supermarket beer purchases, taken over time, were analyzed to describe the strength of company loyalties. Brands were grouped by company with a focus on the three major suppliers—A-B, Miller and Coors—and a remaining category designed to capture brands not supplied by A-B, Miller or Coors—simply referred to as Other. The unconditional average proportion provides information regarding the distribution of beer purchases by company and in turn provides the average market shares by company. Average proportions by company conditional on purchasing at least one brand by a given company are used to describe the strength of company loyalty—i.e. do A-B drinkers remain A-B drinkers. The idea being that company loyalty is high when conditional on purchasing at least one brand supplied by a given company the distribution of beer purchases remains high for that given company. In contrast, company loyalty is low when conditional on purchasing at least one brand supplied by a given company the distribution of beer purchases is higher for any other company. Also, the average proportion of households who only purchased beer brands supplied by a given company describes the prevalence of perfect company loyalty. Lastly, cross-company comparisons establish a measure for a household's willingness to switch and their relative loyalty rankings—households most loyal to one company but least loyal to another.

For the two markets of Eau Claire, WI and Pittsfield, MA. from 2001 until 2006, results in Table 19 indicate that the three major companies—A-B, Miller and Coors—on average account for nearly 70% of supermarket beer sales with Miller leading at 36% followed by A-B and Coors at 26% and 8%, respectively. Conditional average proportions indicate that households who consumed at least one Miller product continued to purchased Miller products with a relative frequency of 74%. Households who purchased at least one A-B product were sightly less loyal with a relative frequency of 66% for A-B brands. For the Other category, households who purchased at least one non A-B, Miller or Coors brand tended to exhibit strong loyalty with a relative frequency of 65% for all other brands. The least loyal were households who purchased at least one Coors brand with a relative frequency of 53% for Coors brands. Conditional on purchasing a brand by a given company, the willingness of households to purchase brands made by a different company occurs with a relative frequency between 11% to 19% when switching to brands by A-B, Miller or all others. The propensity to switch to brands by Coors is rare with a frequency between 3% to 5%. Lastly, perfect company loyalty is surprisingly high as 24%, 18% and 16% of households only purchased brands made by Miller, Other and A-B, respectively, but is weakest for Coors at 4%.

Table 20 continues to analyze household preferences but focuses on package size loyalty. Package categories were grouped by typical sizes: 6-pack, 12-pack, and so on. Two additional sizes were added to capture individual sizes and other smaller sizes smaller than a 6-pack—grouped together as the less than 6-pack category—as well as larger sizes in excess of a 30-pack—grouped together as the more than 30-pack category. The unconditional average proportion provides information regarding the distribution of beer purchases by package size and in turn provides the average market shares by package size. Average proportions by package size conditional on making at least one beer purchase of a given package size are used to describe the strength and degree of package size loyalty. This analysis looks to answer two basic questions. First, do households who purchase beer of a given package size predominately purchase that package size—i.e. do 6-pack drinkers predominately stay 6-pack drinkers? Second, do households who purchase beer of a given package size substitute more towards larger or smaller package sizes—i.e. are 12-pack drinkers more likely to purchase 24-packs more so than 6-packs? Lastly, the average proportion of households who only purchased a particular package size is used to describe the prevalence of perfect package size loyalty.

For the two markets of Eau Claire, WI and Pittsfield, MA from 2001 until 2006, results in Table 20 indicate that package sizes of at least a 12-pack on average account for 75% of beer sales. The more common packages of 6-packs, 12-packs and 24-packs on average account for 21%, 29% and 18% of beer sales. Analyzing conditional means indicates a considerable degree of package loyalty. Specifically, households who purchase at least 1 6-pack tend to stick with 6-packs with a relative frequency of 44%. This pattern strengthens for larger packages. That is, households who purchase at least 1 12-pack tend to purchase 12-packs more frequently than all other sizes with a relative frequency of 58%. For 18-packs, the relative frequency is 47% and for the remaining larger sizes of 24-packs and 30-packs the relative frequency is 58% for each.

In terms of likely deviations from specific package size loyalties, evidence overall suggests a willingness to deviate to larger package sizes. For smaller than 6-pack purchasers, deviations to larger package sizes occur with an equal relative frequency of 16% for beer purchases of both 6-packs and 12-packs. For 6-pack purchasers, deviations to 12-packs occur with a relative frequency of 19% while deviations to the smaller sizes of less than 6-packs occur with a relative frequency of 3.7%. For 12-pack purchasers, the immediate deviation seem to indicate a willingness to deviate to 6-packs, yet taken as a whole larger than a 12-pack are more likely with a relative frequency of 27% compared against the relative frequency of 25% for deviations to package sizes smaller than a 12-pack. For 18-packs, deviations to 24-packs occur with a relative frequency of 20% while deviations to 12-packs occur with a relative frequency of 16%. This relationship reverses when the largest sizes are considered but the relative frequency to smaller package sizes is comparatively low: 12% relative frequency for deviating away from a 24-pack and towards a 18-pack and 10% relative frequency for deviating away from a 30-pack in favor of a 24-pack.

Similar to perfect company preferences, households exhibit non-trivial perfect package size preferences. These tended to be relatively strong for larger package sizes. Specifically, approximately 12%, 16% and 8% of households only purchased 6-packs, 12-packs and 24-packs, respectively. About 5% and 6% of households only purchased 18-packs and 24-packs, respectively. Lastly, 2% of households only purchased the smallest package size category of less than 6-packs.

Table 21 analyzes the average quarterly sales of households who only purchased 6-packs, 12-packs and 24-packs. If these households are drinking at the same rate but shopping at different rates then sales by households who only consumed these package sizes would be proportionally related. That is, the ratio of 24-pack to 6-pack sales would be four and the ratio of 24-pack to 12-pack and 12-pack to 6-pack sales would be two.

However comparing across these households with perfect package preferences for each quarter provides evidence that this proportional relationship does not hold. Specifically, the average sales by quarter suggest households who purchase larger package sizes drink at much higher rates relative to households who purchase smaller package sizes. On a per 12 ounce serving basis across all quarters, average sales suggest households that only purchased 24-packs drank 4 to 5 times more relative to households that only purchased 6-packs and drank 1.25 to 1.60 times more relative to households that only purchased 12-packs. Similarly, households that only purchased 12-packs on a per 12 ounce serving basis drank 2.8 to 3.8 times more relative to households who only purchased 6-packs.

The basic descriptive statistics of available household panel data provide clear trends regarding consumer preferences and behavior that have fundamental implications in estimating the demand for beer. First, the household-level results provide convincing evidence that consumers likely have strong company preferences which may be driven by strong brand preferences. As a result, a consumer's taste for a particular brand should be taken into account when estimating the demand for beer. Second, the household-level results provide convincing evidence that consumers likely have strong package size preferences. Again in a demand estimation context, a consumer's taste for a particular package size should be taken into account. Third and finally, the household-level results provide convincing evidence that consumers who purchase larger package sizes drink more rather than shop less relative to consumers who purchase smaller package sizes. In a demand estimation context, this would provide evidence that consumers are likely not engaging in storing or stockpiling beer when purchasing larger package sizes. Beer also has a short shelf-life becoming stale after about 90 to 110 days and in the short-run is relatively costly to store. Thus, a static model of demand may be a reasonable approach as typical intertemporal issues regarding the storage of durable goods seem unlikely.[15] Additional details concerning the specific demand model, empirical specification and other important determinants

[15]If consumers are storing in response to price discounts then elasticities based on a static discrete choice model will tend to over-estimate the own-price effect and responsiveness. Alternative dynamic models attempt to control for consumer stockpiling, see Hendel and Nevo (2006).

of consumer tastes are discussed in sections (3.5), (3.6) and (3.7).

The household-level data also provide additional insights as to the possible health implications of new beer brands. While company loyalty was shown to be high, it did indicate a willingness to try other brands supplied by other companies. Thus to the extent that the new brands of Michelob Ultra and Bud Select drew existing consumers away from companies, a possible health benefit could occur as existing drinkers would switch to a beer that on the margins features fewer calories and carbohydrates on a per serving basis. However, these changes are small and as a result any potential positive health impact from the initial ultra-light beers from a total calories perspective is also expected to be small.

In contrast, the household-level results on package loyalty, willingness to deviate to larger package sizes and larger package consumers drinking at relatively faster rates suggests that these new brands may lead to serious negative health effects. That is, these new beer brands may not only attract new drinkers but these new drinkers may be more likely to purchase larger package sizes and in turn drink at faster rates. They may also have a willingness to deviate more in favor of larger package sizes. Thus, estimating the extent to which new beer brands attract new drinkers and which package size these new drinkers ultimately choose both are policy-relevant.

3.5 Demand and Supply

This section presents the model of demand and supply used to estimate the extent to which new beer brands attract new drinkers. For demand, consumer behavior is modeled in a static discrete choice framework. That is, a given consumer will choose to purchase one unit of a particular good if and only if the utility from that purchase is at least as large as the utility associated with the purchase of any other good. While initially the presentation follows the more general random coefficients logit demand model, a multinomial logit framework is used due to its computational simplicity. The notation and discussion follows the work of Berry (1994), Nevo (2000, 2001), Knittel et al. (2011) and others.[16]

Assume we observe $t = 1, \ldots, T$ markets, each with $i = 1, \ldots, I_t$ consumers. For each market, we observe aggregate quantities, average prices and product characteristics for J products. Assume the indirect conditional utility of consumer from consuming product in market takes the following quasi-linear form:

$$\begin{gathered} u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt} = V_{ijt} + \varepsilon_{ijt}, \\ i = 1, \ldots, I_t,\ j = 1, \ldots, J,\ t = 1, \ldots, T, \end{gathered} \tag{3.1}$$

where x_{jt} is a K-dimensional row vector of observable non-price characteristics for product j, p_{jt} is the price of product j in market t, ξ_{jt} denotes the product characteristic observed by consumers and firms but not the researcher and ε_{ijt} is a mean-zero stochastic term. Lastly, α_i and β_i are $K + 1$ consumer-specific coefficients and are generally referred to as random coefficients.

The purpose of the random coefficients is to model consumers' tastes. Specifically, it is assumed that consumer-specific tastes depend on a mean value common across all consumers and consumer-specific deviations driven by differences in observable and unobservable consumer characteristics. Formally, this can be compactly re-stated as

[16] These include the general review of demand estimation by Nevo (2011) and Davis et al. (2010) and the specific use of discrete choice models with respect to the U.S. beer industry by Hellerstein (2008) and Goldberg et al. (2010).

follows

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \Pi D_i + \Sigma v_i,\ D_i \sim P_D\left(D\right),\ v_i \sim P_v\left(v\right), \tag{3.2}$$

where D_i is a $d \times 1$ vector of demographic variables, Π is a $(K+1) \times d$ matrix of coefficients that measure how tastes vary with demographics, v_i is a $(K+1) \times 1$ vector which captures unobserved consumer-specific factors and Σ is a $(K+1) \times (K+1)$ matrix of parameters. As an example, consider a consumer that is young and active and as such may have a particularly high valuation for the product characteristic of low calorie beer. The magnitude of this high valuation will depend upon a mean value common across all consumers, the consumer-specific deviation driven by the consumer's age and the consumer-specific deviation driven by the unobservable characteristic of being active.

Since consumers can choose not to consume any of the goods, the demand system is completed by introducing an outside good. The utility of the outside good is normalized to zero. Thus,

$$u_{i0t} = V_{i0t} + \varepsilon_{i0t} = \varepsilon_{i0t},\ V_{i0t} = 0.$$

Define $\theta = (\theta_L,\ \theta_{NL})$ as a vector containing all the parameters of the model where $\theta_L = (\alpha,\ \beta)$ contains all the linear parameters and $\theta_{NL} = (\Pi,\ \Sigma)$ contains all the nonlinear parameters.[17] Next by combining (3.1) and (3.2), the model can be compactly written as the sum of three terms:

$$\begin{aligned} u_{ijt} &= \delta_{jt}\left(x_{jt}, p_{jt}, \xi_{jt}; \theta_L\right) + \mu\left(x_{jt}, p_{jt}, D_i, v_i; \theta_{NL}\right) + \varepsilon_{ijt}, \\ \delta_{jt} &= x_{jt} - \alpha p_{jt} + \xi_{jt},\ \mu_{ijt} = \left[-p_{jt}, x_{jt}\right]^{\intercal} * \left(\Pi D_i + \Sigma v_i\right), \end{aligned} \tag{3.3}$$

where the first term, δ_{jt}, captures the mean utility level common to all consumers associated with the consumption of good j in market t. The sum of the remaining terms, $\mu_{ijt} + \varepsilon_{ijt}$, represents a mean-zero heteroskedastic deviation from a given mean utility level.

Consumers are assumed to purchase one unit of the good that provides the highest utility. In the case of the beer industry, this simplifies to assuming that consumers purchase one package size (i.e. 6, 12, or 24 pack) of a brand of beer that provides the highest utility. Specifically, consumers are defined as a vector of demographics and product-specific shocks, $(D_i, v_i, \varepsilon_{i0t}, \ldots, \varepsilon_{iJt})$, which in turn defines the set of consumer attributes that rationalize the choice of good j. Once this set is defined, it can be used to calculate the market share of good j.

The convention is to define the set of consumers who choose brand in market as follows

[17] This nomenclature exists since as explained by Nevo (2001) the estimation of the full random coefficients logit demand model is estimated by nonlinear GMM. It turns out that the linear parameters can be concentrated out thereby simplifying the computational requirements by having only to search over the nonlinear parameters in order to solve the nonlinear GMM objective function.

$$A_{jt}\left(x_{.t}, p_{.t}, \delta_{.t}; \theta_{NL}\right) = \left\{\left(D_i, v_i, \varepsilon_{i0t}, \dots, \varepsilon_{iJt}\right) | u_{ijt} \geq u_{ilt} \; \forall l = 0, 1, \dots, J\right\},$$

where $x_{.t} = (x_{lt}, \dots, x_{Jt})^{\intercal}$, $p_{.t} = (p_{lt}, \dots, p_{Jt})^{\intercal}$ and $\delta_{.t} = (\delta_{lt}, \dots, \delta_{Jt})^{\intercal}$ are observed product characteristics, prices and mean utilities of all brands, respectively. Assuming ties occur with probability zero, the market share of the j^{th} product is an integral over the mass of consumers in the set A_{jt}. Formally, the market share of the j^{th} product is

$$s_{jt}\left(x_{.t}, p_{.t}, \delta_{.t}; \theta_{NL}\right) = \int_{A_{jt}} dP\left(D, v, \varepsilon\right),$$

which by Bayes' rule and by assuming D_i, v_i and ε are independent simplifies to

$$s_{jt}\left(x_{.t}, p_{.t}, \delta_{.t}; \theta_{NL}\right) = \int_{A_{jt}} dP_{\varepsilon}\left(\varepsilon\right) dP_{v}\left(v\right) dP_{D}\left(D\right), \tag{3.5}$$

where P_{ϵ}, P_v and P_D denote the population distribution of the mean zero stochastic term, unobserved consumer-specific factors and observed consumer-specific factors, respectively. Assumptions on the various disturbance terms yield specific expressions for equation (3.4) which in turn are used to derive specific expressions for calculating the own and cross-price elasticities.[18]

Setting $\theta_{NL} = 0$ thereby setting $\alpha_i = \alpha$ and $\beta_i = \beta$ simplifies the conditional indirect utility shown in equation (3.1) considerably

[18]The advantages and disadvantages of such assumptions have been widely examined by, among others, Berry (1994) and Nevo (2001, 2011).

$$
\begin{gathered}
u_{ijt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \varepsilon_{ijt}, \\
i = 1, \ldots, I_t,\ j = 1, \ldots, J,\ t = 1, \ldots, T.
\end{gathered} \tag{3.6}
$$

Next by making a simplifying assumption regarding ε_{ijt}, we obtain the multinomial logit demand model. Specifically if ε_{ijt} is assumed to be i.i.d. and distributed according to a Type I extreme distribution, then the integral in equation (3.4), and thus the market share of brand j, has a simple closed-form solution

$$
s_{jt} = \frac{exp\left(x_{jt}\beta - \alpha p_{jt} + \xi_{jt}\right)}{1 + \sum_{k=1} exp\left(x_{kt}\beta - \alpha p_{kt} + \xi_{kt}\right)}. \tag{3.7}
$$

The advantage of the logit demand model is its computational simplicity. By an appropriate choice regarding the error term ε_{ijt}, calculating the share of brand j in market t follows a simple closed-form solution. In addition, the elasticities can be easily computed. However, such elasticities in many instances predict unrealistic price and substitution behavior. First, the own-price elasticities tend to be proportional to own price, which would predict that the lower-cost brand has a higher mark-up. Depending on the industry, this may run counter to intuition.[19] Second, the cross-price

[19]For example, in the U.S. auto industry it would seem that the higher priced brands, like BMW, would have the higher mark-up. For the U.S. beer industry, the quality differences across brands are not as large in comparison to the auto industry. Thus, it may be the case that low price beer does experience a higher mark-up. Such would occur if the major suppliers, A-B, Miller and Coors, have exceptionally low marginal cost in producing their lower quality brands. As such, the mark-up for a low quality beer like Milwaukee's Best may in fact be larger in comparison with a higher priced, higher quality import like Guinness. The point is simply that the researcher may want the data to determine if this is indeed the case. However, the logit demand model imposes the relationship that lower-costs brands have higher mark-ups a priori where the source of this a priori restriction, as discussed in Nevo (2000, 2011), is from the functional form assumption regarding how price enters the utility function.

elasticities are only a function of the market share and not product characteristics. As a result, the predicted substitution patterns may be unreasonable.[20] As discussed by Nevo (2000), the source of this limitation is the i.i.d. nature of the error term ε_{ijt} and not its assumed distribution.

Demand estimates can in turn be used with a model of supply in order to estimate marginal cost. These marginal cost values and demand estimates can in turn be used to consider various counterfactuals of interest. The convention in the literature has been to model supply assuming firms behave as Bertrand-Nash competitors; thus, multi-product firms maximize profits by competing in prices. While Bertrand-Nash competition may in fact describe the true nature of competition in an industry, it also allows marginal costs to be recovered in a relatively simple way. As such, the choice of Bertrand-Nash competition may be one motivated by reality or convenience.

A recent study by Rojas (2008) provides an empirically-based justification for assuming firms behave as Bertrand-Nash competitors in the U.S. beer industry. Specifically, Rojas (2008) provides empirical evidence which suggests prices predicted by the assumption that firms compete as Bertrand-Nash competitors match well with the actual pricing behavior of firms in the U.S. beer industry. For this reason, the nature of competition is assumed to be Bertrand-Nash. The specifics of the model are provided below and are primarily based on Nevo (2001), Petrin (2002) and others.[21]

[20] In the case of the U.S. beer industry, suppose we want to examine Sol (an import), Corona Extra (an import) and Bud Light Lime (a domestic) and we assume that Corona Extra and Bud Light Lime have the same market share. Under the logit demand model, if the price of Sol increases, the substitution from Sol towards Corona Extra will be equal to the substitution from Sol towards Bud Light Lime. However one might expect the substitution from Sol towards Corona Extra to be larger since consumers may be more likely to substitute one import for another.

[21] Such would include the treatment in Davis et al. (2010) and the working paper on merger analysis by Knittel et al. (2011).

Suppose there are $\mathcal{F}$ firms, each of which produce some subset, $\Im_f$, of the $j = 1, \ldots, J$ different brands of beer. The profit for firm f is

$$\Pi_f = \sum_{j \in \Im_f} (p_j - mc_j) \, M s_j (p)$$

where $s_j (p)$ is the market share of brand j which is a function of the prices of all other brands, M continues to denote the market size and fixed costs are assumed to be zero. Assuming the existence of a pure-strategy Bertrand-Nash equilibrium in prices and that the prices that support it are strictly positive, the price p_j of any product j by firm f must satisfy the following first-order condition

$$s_j (p) + \sum_{j \in \Im_f} (p_j - mc_j) \frac{\partial s_r (p)}{\partial p_j} = 0 \tag{3.8}$$

To recover the marginal costs, first define the following ownership matrix

$$\Omega^*_{jr} = \begin{cases} 1, & if \; \exists f : \{j, r\} \subset \Im_f, \\ 0, & otherwise, \end{cases}$$

where $j, r = 1, \ldots, J$. Specifically, Ω^*_{jr} acts to select the terms that involve the products produced by firm f. To express the first-order condition in vector notation, define $S_{jr} (p) = -\partial s_r (p) \big/ \partial p_j$ and define the $J \times J$ matrix $\Omega (p)$ as the element-by-element product of Ω^*_{jr} and $S_{jr}(p)$ such that $\Omega_{jr}(p) = \Omega^*_{jr} * S_{jr}(p)$.[22] Then, in vector notation, the first-order conditions can be written as

$$s(p) - \Omega(p)(p - mc) = 0,$$

where $s(\cdot)$, p and mc are $J \times 1$ vectors of market share, price and marginal cost, respectively, and $\Omega(p)$ is a $J \times J$ matrix containing own and cross-price share derivatives. Solving for mc gives a closed-form expression for estimating the marginal costs as prices are observed and demand estimates assist in the recovery of shares and the own and cross-price share derivatives

$$mc = p - \Omega(p)^{-1} s(p). \tag{3.9}$$

Thus equation (3.8) can be used to estimate marginal costs. Once estimated, additional counterfactuals can be considered by predicting a price vector if specific brands were excluded from the choice-set given marginal cost estimates, demand estimates and observed data.

3.6 Data

Grocery store scanner data licensed from Information Resources Incorporated (IRI) is used to estimate demand. The IRI data provide weekly store data for chain

[22] Note, the element-by-element product is sometimes referred in the literature as the Hadamard product, for example see Davis et al. (2010) and Knittel et al. (2011).

grocery stores in 50 markets from 2001 to 2006 where beer brands are tracked at the store-week-UPC level. Each market is a collection of counties centered on a metropolitan area. A natural limitation of the IRI scanner data is that it only applies to grocery stores whereas beer can be sold in a variety of venues including bars, restaurants and convenience stores. Similarly, beer is not necessarily the only alcoholic beverage offered in grocery stores as many may offer wine which may act as a substitute for beer. In addition, grocery stores may be in close proximity to specialty stores which sell liquor which again may act as a substitute for beer. However, industry publications indicate that supermarkets are the leading distributor of beer with a market share of 56.9% whereas specialized retailers account for only 14.4% (Datamonitor 2010). They also indicate that alcohol consumption is dominated by beer with an estimated market share of 84% (Beer Handbook 2011).

Following convention, the IRI was aggregated data up to the city-quarter level for each of the 50 geographic regions for each year. Thus in terms of the demand model discussed in section (3.5), the index for markets, t, indicate a specific city-quarter. Within a market, products are aggregated to the brand-package level. That is, the label name indicates the brand—i.e. Bud Light—which in turn has a variety of available package sizes associated with that brand. Thus, the quantity of units sold for a given brand is then the sum of all units purchased for a given brand-package size across all relevant weeks in a given city-quarter. Similarly, price per unit is the average per unit price for a given brand-package size across all relevant weeks in a given city-quarter.

The IRI data also has information regarding promotions by indicating if during the week a brand-package size was promoted with a minor or major sign or if a price reduction of at least 5% was offered. These three product characteristics were in turn aggregated to the city-quarter level. Hence, they measure the proportion of time a brand-package size was promoted via a minor sign, major sign or price promotion for a given-city quarter. In terms of equation (3.3) in section (3.5), p_{jt} would denote the average price of brand j in city-quarter t for a given package size and the corresponding proportion of times this product was promoted with a minor sign, major sign or price reduction would be included in x_{jt} as a set of observable product characteristic.

To ensure an accurate link between brands and companies, various issues of Adam's Beer Handbook were used in conjunction with the IRI data. Specifically, detailed lists of major suppliers and their major to complete set of beer brands from Adam's Beer Handbook were used to link the brands in the IRI data to companies that produce those brands. A complete listing of top brands in the IRI relative to industry statistics are shown in Table 22. Overall, this approach was able to map all brands to companies which in turn account for on average 96% of the market—defined by the sales in the IRI data from 2001 to 2006. Additionally the popularity of brands according to Adam's Beer Handbook correlated well with the top brands in the IRI data. However, in many cases, the magnitude of IRI average brand shares differed from corresponding industry averages. These differences could be innocuous given the IRI is restricted to 50 markets of beer sales from major grocery stores on average—thus serving consumers with slightly different tastes relative to the industry average. Or it could be a consequence of mapping lesser known brands to companies resulting in a slightly finer segmentation of the brands listed in the IRI relative to the industry.[23]

[23] These would include mapping mult-packs—multiple brands offered in a single pack—as a distinct

Table 23 describes the distribution of sales by package size in the IRI data with high relative frequencies of 38% and 29% for 6-packs and 12-packs, respectively. Smaller to larger package sizes collectively have approximately the same relative frequency with 15% for the category for all sizes less than a 6-pack and 13% for package sizes at least as large as an 18-pack. The largest package sizes of a 24-pack or larger account for about 8% of average sales in the IRI data. When conditioned by company, the distribution of sales by package size by major suppliers—A-B, Miller and Coors—is skewed in favor or larger package sizes relative to the unconditional distribution. The distribution of sales by package size by all other suppliers—all others not including A-B, Miller or Coors—is skewed in favor of smaller package sizes relative to the unconditional distribution. Each are consistent with conventional thought that that macro-brew sales are from larger package sizes and micro-brew sales are from smaller package sizes.

Lastly, Table 24 compares the average proportion of times various promotions occurred in the IRI data from 2001 until 2006 both unconditionally and by company. The types of promotions include a minor sign, a major sign, a price reduction was offered of at least 5%. The variable Any is simply the average proportion if any of the three types of promotions had occurred. By company average proportions are also compared against average company total expenditure shares—company total expenditure on advertising relative to total advertising in the beer industry averaged over time using data from various issues of Adam's Beer Handbook. Unconditional averages indicate price reductions are the dominate promotion with an average frequency of 19%. The by company results also indicate that price reductions are used frequently

brand which were popular and unique to the major companies of A-B, Miller and Coors.

by major companies—A-B, Miller and Coors—but also by all remaining companies with a relative frequency ranging from 21% to 26%. Given the role of advertising in this industry, the average proportion of any of the three types of promotions tracked by the IRI match well in magnitude to average company total expenditure shares. To the extent this relationship continues to hold at the region-brand level, the three types of promotions in the IRI data should act as a proxy for local and national brand advertising.

3.7 Estimation

This section uses the aggregated IRI data based on 50 markets from 2001 until 2006 to estimate demand parameters which in turn are used to evaluate counterfactuals. First, the empirical specification is discussed where remaining model parameters are defined in the context of the U.S. beer industry. Second, identification is discussed with a focus on commonly used instrumental variables in the static demand discrete choice literature. Third, demand estimates are provided. Fourth and lastly, counterfactuals are considered with a focus on evaluating if new brands attract new drinkers and if these new drinkers choose larger package sizes.

As discussed in Berry (1994), the multinomial logit with the utility function as specified by equation (3.5) and corresponding predicted market shares in equation (3.6) can be estimated using the following linear equation:

$$ln\left(s_{jt}\right) - ln\left(s_{0t}\right) = \delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}, \tag{3.10}$$

where s_{jt} denotes the market share of brand j in market t of a given package size and is defined as the ratio of observed units sold of brand j in market t of a given package size relative to the size of the market, denoted M. Following a similar approach used by Cohen (2008), the market size is assumed to equal the maximum of the sum of quarterly sales taken across markets for each quarter and for each year from 2001 until 2006.[24] Next, s_{0t} denotes the share of the outside good—defined as $1 - \sum_j s_{jt}$—and represents the non-alcohol option. For observables, x_{jt} represents observable product characteristics of brand j in market t and includes a set of dummy variables for package size and promotion variables capturing the proportion of times the brand-package size was promoted with any type of sign or with a price reduction of at least 5% in market t.[25] Price per unit for brand j in market t of a given package size is denoted as p_{jt}. Lastly, ξ_{jt} is assumed to represent shifts in demand in response to unobserved promotional activity for brand j in market t.[26] Recall ξ_{jt} appears in the firm's first order condition; thus, p_{jt} and ξ_{jt} are correlated. Identification of the model is achieved by the use of various instruments.

A set of instruments are required to identify the model in equation (3.9). Valid instruments must be correlated with unit prices but uncorrelated with unobserved

[24] Note, Cohen (2008) estimated the demand for paper towels and allowed for multiple package sizes in a random coefficients framework examining evidence of second-degree price discrimination. He used this approach to define a saturation point for total paper towel sales. For clarity in its implementation in this paper, the total number of units sold across all brands and across all cities for each quarter for each year was first calculated. The maximum of this set of totals determined the size of the market. It is tempting to first calculate the total number of units sold across all brands by each city for each quarter for each year and then take the maximum of this alternative set of totals in order to define the size of the market. However, this alternative approach would necessarily lead to a market share of zero for the outside good for the maximal market.

[25] These represent the collection of observed product characteristics, both to the consumers and econometrician. The set of dummy variables for package size include indicators for 6-packs, 12-packs, 18-packs, 24-packs and 30-packs. Larger than 30-packs were dropped and thus the base groups is less than 6-packs.

[26] That is, known and common to consumers but unknown to the econometrician.

demand shocks. That is, to identify demand, candidates for valid instruments will be variables that impact supply—and in turn impact per unit prices—but will not impact demand—will be uncorrelated with the unobserved demand shock ξ_{jt}. Following the literature, three sets of instruments are considered.

The first set of instruments are functions of observable product characteristics following BLP (1995). In the current application, these include the sum of all brands produced by a company in a given market, the sum of all brands produced by all rivals in a given market, and the sum of the proportions of times a brand was promoted—including separately by any type of sign and with a price reduction—by a company in a given market and similarly by all rivals in a given market.[27] These variables are valid instruments so long as product characteristics are assumed to be exogenous.[28] They are also correlated with price from the firm's first order condition since price depends on own-firm and rival-firm observable product characteristics.[29]

The second set of instruments follows from Hausman (1997) and Nevo (2001) and utilizes the panel structure of the data. Specifically, brand fixed-effects are added as additional regressors thereby controlling for consumers time-invariant taste for a particular brand. The unobserved demand shock then represents market-specific deviations from these tastes for a particular brand—that is, $\Delta\xi_{jt} = \xi_{jt} - \xi_j$. In the present context, this deviation would represent a city-specific demand response to unobserved brand advertising separate from a time-invariant taste for that brand. Assuming these deviations are independent across cities but correlated within a city, the average price of a brand taken over other cities can instrument for the price of that

[27] This approach resulted in 6 total instruments.

[28] Or as explained in Nevo (2000), the product characteristics are pre-determined prior to the revelation of the consumer's evaluation of the unobserved product characteristic ξ_{jt}.

[29] See equation (3.6) along with the firm's FOC in equation (3.7).

brand in a given city. For a given brand, these prices are correlated because they share a common marginal cost term in the firm's FOC. However, in the present context, if promotional activity is coordinated across cities, then this approach is invalid and price will remain endogenous. Following this approach, brand fixed effects and year fixed effects are included as additional regressors and average quarterly prices are included as instruments.[30] Year fixed effects are added to further control for cross-market coordination but is clearly imperfect given the role of national, and likely regionally-coordinated, advertising in the beer industry.

The third set of instruments follows Cohen (2008) who estimated demand for paper towels allowing for different package sizes as an additional set of product characteristics. In addition to using functions of observable product characteristics, Cohen (2008) used year dummies as exclusion restrictions in estimating the supply equation believing that year dummy variables proxy for cost shifters under the assumption that costs change over time but preferences for the outside good do not. Thus, in the present context, functions of product characteristics will continue to be used as instruments but year dummy variables will also be considered.[31] The model is initially estimated by ordinary least squares to evaluate the impact on the estimate of the price coefficient as various fixed effects are added including year, brand and city. Next to control for the endogeneity of price, the model is estimated by two-stage least squares where each of the three aforementioned sets of instruments are evaluated across two specifications involving additional regressors: the first includes brand

[30] Following Nevo (2001), quarterly average prices across cities minus the city being instrumented were added as instruments. This resulted in 24 total instruments deriving from the fact there are 4 quarters in each of the 6 years of IRI data.

[31] This approach results in 11 total instruments—6 from the product characteristics and 5 for the year dummies with 2001 serving as the base year.

and year fixed-effects and the second includes brand, year and city fixed-effects.[32]

Logit demand parameters are initially estimated by OLS with the results shown in Table 25 where each column adds an additional set of fixed-effects. Across all OLS specifications, mean utility increases from larger package sizes. Also, mean utility increases as the proportion of times a brand is advertised with a minor or major sign increases. Similarly, mean utility increases as the proportion of times a brand is advertised with a price reduction of at least 5%. Given these results, it seems reasonable to suppose that the correlation between unobserved advertising and mean utility is positive. In addition, assuming that unobserved advertising increases the cost to produce a given brand, the correlation between unobserved advertising and price is likely positive. Combined, this implies that the estimated coefficient for price shown in Table 25 are likely biased towards zero. Further, the own-price elasticity estimates will be biased towards inelastic demand for a given brand-package size, ceteris paribus. Properly controlling for this endogeneity should increase the estimate on price in absolute value.

Logit demand parameters estimated by two-staged least squares are provided in Table 26. The first two columns uses average prices in other markets as instruments. While the first-stage results suggest these instruments are strong, the coefficient on price remains low in absolute value and essentially unchanged relative to the ordinary least squares estimates in Table 25. This would suggest the bias from unobserved promotions likely remains due to coordination in unobserved promotion activity across cities. Using functions of product characteristics as instruments in columns (3) and (4), the price coefficient increased dramatically in absolute value and consumer's

[32]The exception is the specification motivated by Cohen (2008) when year fixed-effects are used as instruments; thus, they are excluded from the demand equation.

display a strong taste for larger package sizes and a continued taste for advertised brands.[33] However, first-stage F-tests were the weakest.[34] Lastly adding both functions of product characteristics and year dummies—proxies for input prices, the first stage F-tests improve but the price coefficient drops dramatically in absolute value indicating that price remains endogenous. This would occur if changes in costs over time were in part due to changes in the use of unobserved promotions over time which in turn introduces correlation between price and unobserved promotions.

For counterfactuals designed to estimate the extent to which the initial ultra-light beers of Michelob Light and Bud Select attracted new drinkers, the demand estimates from column (4) in Table 26 will be used. While statistically weak relative to the remaining specifications, they do reflect a strong preference for package-size—which was implied by the IRI household-level data—and would predict more elastic own-price estimates relative to the remaining specifications.[35] The statistical performance of this preferred specification could possibly improve with the use of input prices as additional instruments and by controlling for market-specific consumer demographics.

Using demand estimates from column (4) in Table 26, counterfactuals are used to evaluate the extent to which Michelob Ultra and Bud Select attracted new drinkers. These counterfactuals are also used to determine if new drinkers are more likely to

[33]Notice also in column (3) relative to column (4), omitting possible city-specific tastes after controlling for the endogeneity in price tended to understate consumer tastes for larger package sizes.

[34]Further, the variance-covariance matrix for the first stage regression in column (3) and similarly for column (5) was almost singular and thus reported no standard errors. Similar singularity issues also occurred in BLP (1995).

[35]That is, the own-price elasticity estimates may very well be inelastic for most brands using the BLP specification; but given the higher price estimate in absolute value, they would be more elastic relative to the remaining specifications. From principles, given the current product scope, own-price elasticities should be elastic given the ample number of possible close-substitutes. Of course, brand-loyalties could be strong enough to overcome this effect.

purchase larger package sizes. In the context of the demand model, the thought experiment is to measure the number of consumers in a given market that choose Michelob Ultra (and similarly for Bud Select) when it is available in the choice-set. Additionally, the package size choice can also be determined. After identifying this set of consumers, counterfactuals involve removing these brands from the choice set and tracking how many of these former purchasers of Michelob Ultra (and similarly for Bud Select) switched to the outside good—all other goods except beer. Given the data consists of grocery store beer purchases, the outside good serves as the non-alcohol option.[36] Hence, analyzing how many former purchasers switch to the outside good once the brand is removed acts as the measure of how strong a new brand attracted new drinkers.

For simplicity, a single market was analyzed in 2003 for Michelob Ultra and 2005 for Bud Select.[37] The general procedure for each of these two new brands worked as follows. For each year and new brand, it was assumed each market had 50,000 consumers. To estimate the number of times the new brand was chosen, a vector of mean utilities, common to all consumers in the case of the logit model, was constructed by predicting values from equation (3.9) using estimated demand parameters with the last element of this vector set to zero representing the normalized mean utility of the outside good. To estimate the utility of each good including the outside good, each consumer has a market-brand-package size specific i.i.d. taste shock drawn from a type 1 extreme value distribution which was added to predicted mean utility.

[36] While not ideal, any substitution from these initial ultra-light beers to grocery store offered wine is suspected to be small.

[37] This was the first market in each of the two selected years which in turn were the first full year for each new brand. Additionally, it was relatively small thus was very tractable. An alternative and conventional approach would have been to estimate these counterfactuals for all markets in 2003 for Michelob Ultra and similarly for 2005 for Bud Select.

Consumers then choose across all available brands in the given market the product that yields the highest utility.

To estimate the counterfactual product choices when these new brands were removed from the choice set, the following approach was used. First, using demand estimates and observable data on prices and product shares, marginal costs for each product were recovered using equation (3.8). Second, the new brands for the single market considered for the respective years were removed from the choice set and counterfactual prices were estimated by solving each firm's first-order equation:

$$s\left(p^{cf}\right) - \Omega\left(p^{cf}\right)\left(p^{cf} - \hat{mc}\right) = 0,$$

where predicted market shares, $s\left(p^{cf}\right)$, are defined by equation (3.6), evaluated at mean utilities net of price and demand parameters and are a function of counterfactual prices. The ownership matrix, $\Omega\left(p^{cf}\right)$, is similarly defined as it involves derivatives of the predicted market share.[38] Lastly, $\hat{mc}$ denotes the vector of recovered marginal costs. Once counterfactual prices were recovered, predicted utilities were obtained using the same aforementioned procedure.[39] The counterfactual results are provided in Table 27.

For a single market in 2003, the estimated number of consumers who chose Michelob Ultra when it was in the choice-set was 222 with 156 choosing 6-packs and 66 choosing 12-packs. When Michelob Ultra was excluded, 165 consumers switched to the outside good. Thus, for this single market, 74% of Michelob Ultra sales were

[38] However, any changes in conduct as a result of mergers were ignored for simplicity.

[39] Note, the only difference is the new counterfactual prices and the reduced choice-set.

generated by new drinkers and 26% were generated by existing drinkers.

For the same single market but in 2005, the estimated number of consumers who chose Bud Select when available in the choice-set was 69. These consumers favored larger package sizes with 60 choosing 12-packs and 9 choosing 6-packs. When Bud Select was excluded, 47 consumers switched to the outside good. Thus, for this single market, 68% of Bud Select sales were generated by new drinkers and 32% were generated by existing drinkers.

3.8 Conclusion

Counterfactual results suggest that sales of the initial ultra-light beers of Michelob Ultra and Bud Select were predominately generated by new drinkers. Based on a single market, 68% of sales of Bud Select and 74% of sales of Michelob Ultra were due to new drinkers. Additionally, new drinkers of Bud Select seemed to prefer larger package sizes. Given the various public health concerns regarding alcohol consumption, these results appear alarming. However there are many notable limitations. First, the logit demand model does impose substantial restrictions on the own- and cross-price effects, firm mark-ups and mean utilities. Second, the counterfactuals considered were only for one market for each new brand and ignored any changes in brand-ownership as a result of mergers for simplicity.

Future work should analyze the sensitivity of the results with adjustments in the size of the market as the current brand-package size shares seemed too small. An alternative saturation-point should be considered as well as other approaches used in the literature—including a per-serving approach and an average number of shopping trips approach. Other instruments should be considered including input prices. For additional variation, regional input prices like gasoline and water by week or by quarter may be useful. By year mappings from brands to companies could be developed to incorporate changes in brand-ownership as a result of mergers. Counterfactuals could be extended to all markets for the initial year a new brand was available to obtain a complete distribution in terms of its effectiveness to attract new drinkers. Lastly, a more flexible demand model like the random-coefficients logit model should be considered.

3.9 Tables

Table 16. Light Beer Sales from 2001 to 2006.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Bud Light	*Miller Light*	*Coors Light*	*Michelob Light*	*Michelob Ultra*	*Bud Select*	*Michelob Ultra Amber*
By Year:							
2001	469.5	217	231	40			
2002	505	214.5	232.5	40.4	5.9		
2003	517	217	228.95	32	41.5		
2004	536	241	224.37	26	58		
2005	536.7	248.2	228	23	45	32.4	.1
2006	560	246.7	231.7	21	42	30	6
Company:	*A-B*	*Miller*	*Coors*	*A-B*	*A-B*	*A-B*	*A-B*

Source: Adams Beer Handbook (2007).
Notes: Units are millions of 2.25-gallon cases sold. *A-B* denotes Anheuser-Busch Brewing Company; *Miller* denotes Miller Brewing Company; and, *Coors* denotes Coors Brewing Company.

Table 17. Advertising Expenditures as a Share of Company and Industry Total Advertising Expenditure by Select Brands.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Bud Light	*Miller Light*	*Coors Light*	*Michelob Light*	*Michelob Ultra*	*Bud Select*	*Michelob Ultra Amber*
By Year:							
2001	28	43	60	12			
2002	27	43	66	11	4.4		
2003	32	53	74	4.2	11		
2004	34	46	74	0.37	14		
2005	30	70	87	0.43	8	16	0
2006	37	58	86	0.06	6.1	21	4
By Industry:							
Beer	12	11	12	0.89	3.5	7.6	1.5
Alcohol	8.5	7.3	8	0.61	2.4	5	0.97
Company:	*A-B*	*Miller*	*Coors*	*A-B*	*A-B*	*A-B*	*A-B*

Source: Adams Beer Handbook various issues.
Notes: By year brand-specific shares are expressed as a percentage relative to company total advertising expenditure. By industry shares are median values taken over time—2001 to 2006—and are expressed as a percentage relative to the total advertising expenditure of the entire beer and alcohol industry. Lastly, *A-B* denotes Anheuser-Busch Brewing Company; *Miller* denotes Miller Brewing Company; and, *Coors* denotes Coors Brewing Company.

Table 18. Characteristics of Select Light and Ultra-Light Beers.

	(1)	(2)	(3)
	Calories	*Carbs.*	*A.B.V.*
A-B:			
Bud Light	110	6.6g	4.2%
Michelob Ultra	95	2.6g	4.2%
Bud Select	99	3.1g	4.3%
Bud Select 55	55	1.9g	2.4%
Miller:			
Miller Lite	96	3.2g	4.2%
MGD 64	64	2.4g	2.8%
Coors:			
Coors Light	104	5.3g	4.15%
Aspen Edge	94	2.6g	4.1%

Source: www.beer100.com

Notes: *A-B* denotes Anheuser-Busch Brewing Company; *Miller* denotes Miller Brewing Company; and, *Coors* denotes Coors Brewing Company. Calories, grams of carbohydrates (*Carbs.*) and percent alcohol by volume (*A.B.V.*) are based on a per 12 ounce serving.

Table 19. Sample Average of Household Company Preferences.

	(1)	(2)	(3)	(4)
	A-B	*Miller*	*Coors*	*Other*
Unconditional Average:	26.1	36.4	8.0	29.6
Conditional Averages:				
Purchased at least 1 A-B Product	66.3	16.4	4.1	13.2
Purchased at least 1 Miller Product	11.2	73.6	2.7	12.5
Purchased at least 1 Coors Product	14.5	12.5	53.3	19.7
Purchased at least 1 Other Product	12.2	17.4	5.0	65.4
Perfect Loyalty:	15.8	23.5	4.2	18.2

Source: IRI Household Panel Data of two Behavior Scan Markets: Eau Claire, Wisconsin and Pittsfield, Massachusetts.

Notes: Each cell represents an indicated average taken over time—2001 to 2006. *A-B* denotes Anheuser-Busch Brewing Company; *Miller* denotes Miller Brewing Company; and, *Coors* denotes Coors Brewing Company. The category Other includes all other companies excluding A-B, Miller and Coors. *Unconditional Average* gives the average distribution of purchases by households across companies. *Conditional Averages* gives the average distribution of purchases across companies by households who purchased at least one product from a given company. *Perfect Loyalty* denotes the average percentage of households who only purchased brands made by A-B, Miller, Coors or Other.

Table 20. Sample Average of Household Package Preferences.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Less 6p	*6p*	*12p*	*18p*	*24p*	*30p*	*More 30p*
Unconditional Average:	4.9	20.6	29.3	11.8	17.5	12.5	0.018
Conditional Averages:							
Smaller than 6 pack	44.2	15.6	15.8	6.3	8.6	5.4	0.008
6 pack	3.7	58.4	18.6	5.0	7.2	5.0	0.009
12 pack	2.4	12.2	59.9	7.1	9.6	6.7	0.016
18 pack	2.0	6.8	15.6	46.5	19.7	7.2	0.015
24 pack	1.9	6.2	13.5	12.1	57.7	7.0	0.006
30 pack	1.7	6.3	14.3	7.1	10.4	58.2	0.028
Perfect Loyalty:	2.4	11.7	15.8	4.9	7.7	6.3	0.007

Source: IRI Household Panel Data of two Behavior Scan Markets: Eau Claire, Wisconsin and Pittsfield, Massachusetts.

Notes: Each cell represents an indicated average taken over time—2001 to 2006. *Unconditional Average* gives the average distribution of purchases by households across package size. *Conditional Averages* gives the average distribution of purchases across package size by households who purchased at least one product of a given package size. *Perfect Loyalty* denotes the average percentage of households who only purchased one indicated package size.

Table 21. Average Total Packages Purchased by Households with Strict Package Preferences.

	(1)	(2)	(3)	(4)
	Q1	*Q2*	*Q3*	*Q4*
Only Purchased:				
6 pack	97	140	171	143
12 pack	152	251	244	254
24 pack	98	180	199	157

Source: IRI Household Panel Data of two Behavior Scan Markets: Eau Claire, Wisconsin and Pittsfield, Massachusetts.

Notes: 6 pack denotes households who only consumed 6 packs; 12 pack denotes households who only consumed 12 packs; 24 pack denotes households who only consumed 24 packs. Averages of each quarter for each type of household were taken over time—2001 to 2006.

Table 22. Average Annual Market Share by Brand.

	(1)	(2)	(3)
	Avg. IRI Share	*Avg. Adams Share*	*Avg. Adams Rank*
Top Brands in IRI:			
BUDWEISER	5.5	13.8	2
BUD LIGHT	5.4	18.4	1
MILLER LITE	4.7	8.3	3
COORS LIGHT	4.4	8.1	4
MILLER GENUINE DRAFT	3.3	2.0	11
HEINEKEN	2.7	2.2	10
CORONA	2.4	3.5	6
COORS	2.2	0.7	22
SMIRNOFF TWISTED V	2.0	0.3	19
CORONA LIGHT	1.8	0.3	40
MILLER HIGH LIFE	1.6	2.5	9
MICHELOB ULTRA	1.4	1.2	8
SMIRNOFF ICE	1.2	0.3	22
MICHELOB LIGHT	1.2	1.0	16
Unaccounted Brands:			
All Others	4.3		

Sources: IRI (50 markets); Adams Beer Handbook various issues.
Notes: All averages were taken over time—2001 to 2006. All shares are expressed as a percent. *Unaccounted Brands* denotes the total average share across all brands that were not specifically mapped to a company using industry publications. See the text for further details.

Table 23. Sample Average of Package Size by Company.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Less 6p	*6p*	*12p*	*18p*	*24p*	*30p*	*More 30p*
Full Sample:	15.1	37.8	29.2	4.8	5.1	2.8	0.07
By Company:							
A-B	10.9	34.2	30.8	7.2	7.1	3.0	0.10
Miller	16.1	22.8	34.6	8.1	8.3	5.2	0.07
Coors	8.7	33.5	30.7	9.3	7.4	5.0	0.10
Other	18.8	48.2	25.4	0.7	1.7	1.0	0.04

Source: IRI (50 markets).

Notes: Each cell represents an indicated average taken over time—2001 to 2006. *Full Sample* gives the average distribution of purchases by package size. *By Company* conditions this distribution on purchases by company where *A-B* denotes Anheuser-Busch Brewing Company; *Miller* denotes Miller Brewing Company; and, *Coors* denotes Coors Brewing Company. Lastly, the category Other includes all other companies excluding A-B, Miller and Coors.

Table 24. Sample Average of Promotions by Company.

	(1)	(2)	(3)	(4)	(5)
	Minor	*Major*	*Price*	*Any*	*Adams*
Full Sample:	5.5	3.1	19.2	23.5	
By Company:					
A-B	6.3	3.6	15.9	21.2	38.0
Miller	5.7	3.0	16.4	20.9	21.8
Coors	6.0	3.5	19.3	23.7	16.0
Other	4.7	2.6	22.6	26.0	24.2

Sources: IRI (50 markets); Adams Beer Handbook various issues.
Notes: Each cell represents an indicated average taken over time—2001 to 2006. *Full Sample* gives the average frequencies of various store-promotions including: a minor display (*Minor*), a major display (*Major*) a price reduction of at least 5% (*Price*) and any promotion (*Minor*, *Major* or *Price*). *By Company* conditions by company where *A-B* denotes Anheuser-Busch Brewing Company; *Miller* denotes Miller Brewing Company; and, *Coors* denotes Coors Brewing Company. The category Other includes all other companies excluding A-B, Miller and Coors. Lastly, Column 5 shows the average share of total advertising expenditure by company relative to the total advertising expenditure of the beer industry from 2001 until 2006 based on data in the Adams Beer Handbook.

Table 25. OLS Logit Demand Estimates.

	(1)	(2)	(3)	(4)
	Basic	*Add: Year*	*Add: Brand*	*Add: City*
Price	-0.173***	-0.173***	-0.138***	-0.127***
	(0.018)	(0.018)	(0.027)	(0.022)
6 pack	0.351***	0.352***	1.162***	1.295***
	(0.094)	(0.094)	(0.108)	(0.077)
12 pack	1.362***	1.365***	1.422***	1.504***
	(0.120)	(0.122)	(0.201)	(0.161)
18 pack	1.465***	1.468***	0.598**	0.607**
	(0.153)	(0.155)	(0.271)	(0.227)
24 pack	1.627***	1.632***	0.992***	0.965***
	(0.212)	(0.216)	(0.353)	(0.296)
30 pack	1.850***	1.854***	1.399***	1.486***
	(0.189)	(0.192)	(0.392)	(0.347)
Any Display	2.794***	2.793***	1.784***	1.996***
	(0.278)	(0.277)	(0.224)	(0.150)
Price Break	1.546***	1.547***	1.692***	1.241***
	(0.191)	(0.193)	(0.172)	(0.102)
Fixed Effects:				
Year	N	Y	Y	Y
Brand	N	N	Y	Y
City	N	N	N	Y
Observations	378,310	378,310	378,310	378,310
R-squared	0.10	0.10	0.43	0.50

Notes: White-robust standard errors clustered by city shown in parentheses. For all columns the dependent variable is the difference in log shares. Levels of statistical significance are indicated as follows: *10%; **5%; ***1%.

Table 26. IV Logit Demand Estimates.

	A. Nevo		*B. BLP*		*C. Cohen*	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Price	-0.164***	-0.155***	-1.324***	-2.792***	-0.770***	-0.427***
	(0.027)	(0.024)	(0.454)	(0.943)	(0.172)	(0.066)
6 pack	1.023***	1.166***	5.484***	10.878***	3.464***	2.372***
	(0.105)	(0.092)	(1.698)	(3.344)	(0.662)	(0.263)
12 pack	1.521***	1.639***	10.893***	22.663***	6.469***	3.884***
	(0.205)	(0.186)	(3.667)	(7.426)	(1.400)	(0.537)
18 pack	0.692***	0.746***	13.021***	28.398***	7.222***	3.734***
	(0.263)	(0.242)	(4.821)	(9.792)	(1.843)	(0.708)
24 pack	1.315***	1.352***	16.725***	36.259***	9.373***	4.934***
	(0.368)	(0.343)	(6.102)	(12.575)	(2.331)	(0.901)
30 pack	1.676***	1.835***	18.563***	39.834***	10.560***	5.804***
	(0.398)	(0.358)	(6.631)	(13.575)	(2.524)	(0.964)
Any Display	1.977***	2.227***	1.806***	2.542***	1.787***	2.054***
	(0.261)	(0.192)	(0.281)	(0.413)	(0.232)	(0.158)
Price Break	1.728***	1.128***	1.501***	0.641*	1.605***	1.174***
	(0.194)	(0.100)	(0.249)	(0.331)	(0.187)	(0.097)
Fixed Effects:						
Year	Y	Y	Y	Y	N	N
Brand	Y	Y	Y	Y	Y	Y
City	N	Y	N	Y	N	Y
Observations	284,892	284,892	378,310	378,310	378,310	378,310
First Stage:						
R-squared	.95	.96	.90	.91	.90	.91
F-test	818	1120	9.3	2.3	49	48

Notes: White-robust standard errors clustered by city shown in parentheses. For all columns the dependent variable is the difference in log shares. Levels of statistical significance are indicated as follows: *10%; **5%; ***1%.

Table 27. Counterfactual Results.

	(1)	(2)
	Michelob Ultra	*Bud Select*
A. Brand Included:		
N_j	222	69
By Package:		
6-pack	156	9
12-pack	66	60
B. Brand Excluded:		
N_0	165	47
C. Generated Sales:		
New Drinkers	0.74	0.68
Existing Drinkers	0.26	0.32

Notes: Counterfactual results for a single market for Michelob Ultra in 2003 and again for Bud Select in 2005. For each year and new brand, the number of consumers was set to 50,000. The number of consumers who chose an indicated brand is denoted N_j. The number of consumers who chose an indicated brand when it was in the choice-set but when excluded switched to the outside good is denoted N_0. For further details, see the text.

REFERENCES

[1] Ackerberg, D. A. and M. Rysman. 2005. "Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects." The RAND Journal of Economics, Vol. 36, No.4, pp. 771-788.

[2] Ackerberg, D. A., C.L. Benkard, S. Berry and A. Pakes. 2007. "Econometric Tools for Analyzing Market Outcomes." Handbook of Econometrics, Vol. 6A.

[3] Adams Beer Handbook. The Beverage Information Group: Various Issues: 2001-2007, 2011.

[4] Alexander, Donald, William Kern and Jon Neill "Valuing the Consumption Benefits from Professional Sports Franchises." Urban Economics 48, 321-337 (2000).

[5] American League Red Book. The Sporting News, annual issues: 1986-2009.

[6] American Lung Association, Clean Air Task Force and Earthjustice. "Sick of Soot: How the EPA can save lives by cleaning up fine particle air pollution." (2011).

[7] Anderson, Michael L. "The Benefits of College Athletic Success: An Application of the Propensity Score Design with Instrumental Variables." NBER Working Paper No. 18196 (June, 2012).

[8] Anderson, Jonathan O., Josef G. Thundiyil and Andrew Stolbach. "Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health." Journal of Medical Toxicology, Volume 8, Issue 2, (June, 2012), pp 166-175.

[9] Associated Press. "Senate Votes Yes on Vikes' Stadium." Accessed via ESPN, (May, 2012).

[10] Baade, Robert, "Is There an Economic Rationale for Subsidizing Sports Stadiums?" The Heartland Institute, Policy Study No. 13 (February, 1987).

[11] Baade, Robert, "Stadiums, Professional Sports, and Economic Development: Assessing the Reality," Heartland Institute Policy Study, April 4 (1994).

[12] Baade, Robert and Allen Sanderson, “The Employment Effect of Teams and Sports Facilities.” Sports, Jobs and Taxes, R. Noll, A. Zimbalist (Eds.), Brookings Institution Press, Washington, DC, 1997, pp. 92 – 118.

[13] Baade, Robert A. “The Economic Impact of Mega-Sporting Events.” Handbook of the Economics of Sport, Wladimir Andreff and Stefan Szymanski (Eds.), Edward Elgar Publishing (2006).

[14] Baltagi, Badi H., Econometric Analysis of Panel Data. 4th Edition, Wiley, 2008.

[15] Berry, S. 1994. “Estimating Discrete-Choice Models of Product Differentiation.” The RAND Journal of Economics, Vol. 25, No. 2, pp. 242-262.

[16] Berry, S., J. Levinsohn and A. Pakes. 1995. “Automobile Prices in Market Equilibrium.” Econometrica, 841-90.

[17] Bevis, Charlie, “Holiday Doubleheaders.” The Baseball Research Journal 33, 60-63 (2004).

[18] Bevis, Charlie, Double-Headers: A Major League History. McFarland, 2011.

[19] Boyd, Thomas C. and Timothy C. Krehbiel, “An Analysis of the Effects of Specific Promotion Types on Attendance at Major League Baseball Games.” Mid-American Journal Of Business, Vol. 21, No. 2 (2006).

[20] Bray, J.W., B. Loomis and M. Engelen. 2007. “Correlates of in-store promotions for beer: Differential effects of market and product characteristics.” Journal of Studies on Alcohol and Drugs, 68 (2): 220-227.

[21] Bray, J.W., B.R. Loomis, and M. Engelen. 2009. “You save money when you buy in bulk: Does volume-based pricing cause people to buy more beer?” Health Economics, 18 (5): 607-618.

[22] Bresnahan, Timothy F., and Peter C. Reiss. “Entry in Monopoly Markets.” Review of Economic Studies. (1990) 57, 531-553.

[23] Bresnahan, Timothy F., and Peter C. Reiss. “Entry and Competition in Concentrated Markets.” Journal of Political Economy, Vol. 99, No. 5. (October, 1991), pp. 977-1009.

[24] Bronnenberg, Bart J., Michael W. Kruger, Carl F. Mela. 2008. “Database paper: The IRI marketing data set.” Marketing Science, 27 (4) 745-748.

[25] Brook, Robert D., Sanjay Rajagopalan, C. Arden Pope III, Jeffrey R. Brook, Aruni Bhatnagar, Ana V. Diez-Roux, Fernando Holguin, Yuling Hong, Russell V. Luepker, Murray A. Mittleman, Annette Peters, David Siscovick, Sidney C. Smith, Jr, Laurie Whitsel and Joel D. Kaufman. "Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement from the American Heart Association." Circulation: Journal of the American Heart Association, (2010).

[26] Browning, Robert A. and Louisa S. DeBolt, "The Effects of Promotions on Attendance in Professional Baseball." Sport Journal, Vol. 10, No. 3 (2007).

[27] Bruggink, Thomas H. and James W. Eaton, "Rebuilding Attendance in Major League Baseball: The Demand for Individual Games." Baseball Economics: Current Research, John Fizel, Elizabeth Gustafson, and Lawrence Hadley (Eds.), 9-31, (1996). Wesport, CT: Praeger.

[28] Cameron, Colin A. and Pravin K. Trivedi, Microeconometrics Using Stata. Stata Press, 2009.

[29] Carlino, Jerry and Edward Coulson, "Should Cities Be Ready for Some Football? Assessing the Social Benefits of Hosting an NFL Team." Business Review of the Federal Reserve Bank of Philadelphia, Q2, 2004, 7-17.

[30] Carlino, Jerry and Edward Coulson, "Compensating Differentials and the Social Value of NFL Franchises" Journal of Urban Economics, July 2004, 56, 25-50.

[31] Carlino, Jerry and Edward Coulson, "Compensating Differentials and the Social Value of NFL Franchises: Reply." Journal of Urban Economics, July 2006, 60, 132-138.

[32] Carpenter, C. and C. Dobkin. 2010. "Alcohol Regulation and Crime." NBER Working Paper 15828.

[33] Clapp, Christopher M. and Jahn K. Hakes, "How Long a Honeymoon? The Effect of New Stadium on Attendance in Major League Baseball." Journal of Sports Economics, Vol. 6, No. 3 (2005).

[34] Coates, Dennis, and Brad Humphreys, "The Growth Effects of Sports Franchises, Stadia and Arenas. Journal of Policy Analysis and Management 18 (4), 601-624, 1999.

[35] Coates, Dennis and Brad R. Humphreys, "The Effect of Professional Sports on Earnings and Employment in the Services and Retail sectors in US cities." Regional Science and Urban Economics 33 (2003) 175-198.

[36] Coates, Dennis, Brad Humphreys and Andrew Zimbalist, "Compensating differentials and the social benefits of the NFL: A comment." Journal of Urban Economics 60 (2006) 124-131.

[37] Coates, Dennis and Craig A. Depken II, "The Impact of College Football Games on Local Sale Tax Revenue: Evidence from Four Cities in Texas." Eastern Economic Journal, 2009, 35, (531–547).

[38] Coates, Dennis and Craig A. Depken II, "Mega-events: Is Baylor Football to Waco what the Super Bowl is to Houston?" Journal of Sports Economics, Dec 2010.

[39] Daughters, Amy. "College Football: Should FBS vs. FCS Games be Banned Permanently?" Bleacher Report, (September, 2010).

[40] Davis, P. and E. Garces. 2010. Quantitative Techniques for Competition and Antitrust Analysis. Princeton University Press.

[41] Deaton, A. and J. Muellbauer. 1980a. "An Almost Ideal Demand System." American Economics Review, Vol. 70, No. 3, pp. 312-326.

[42] Deaton, A. and J. Muellbauer. 1980b. Economics and Consumer Behavior. Cambridge University Press, pp. 60-82.

[43] Fixmer, Andy. "Los Angeles City Council Debates AEG's NFL Stadium Plans." Bloomberg Businessweek: News From Bloomberg, (September, 2012).

[44] Gentzkow, M. 2007. "Valuing New goods in a Model with Complementarity: Online Newspapers." American Economic Review, Vol. 97, No. 3, pp. 713-744.

[45] Glassman, Tavis, Chudley E. Werch, Edessa Jobli, and Hui Bian, "Alcohol-Related Fan Behavior on College Football Game Day." Journal of American College Health (2007), Vol. 56, No. 3.

[46] Goldberg, P. and R. Hellerstein. 2010. "A Structural Approach to Identifying the Sources of Local-Currency Price Stability." Working paper: under review for the Review of Economic Studies.

[47] Hansen, B. 2013. "Punishment and Deterrence: Evidence from Drunk Driving." Working Paper.

[48] Hausman, J., G. Leonard and J.D. Zona. 1994. "Competitive Analysis with Differentiated Products." Annales D'Economie Et De Statistique No. 34.

[49] Hausman, J. 1997. "Valuation of New Goods Under Perfect and Imperfect Competition." The Economics of New Goods, Studies in Income and Wealth Vol. 58, ed. By T. Bresnahan and R. Gordon. Chicago: NBER.

[50] Hausman, J. 1999. "Cellular Telephone, New Products and the CPI." Journal of Business and Economic Statistics. Vol. 12, No. 2: pp. 188-194.

[51] Hellerstein, R. 2008. "Who Bears the Cost of a Change in the Exchange Rate? Pass-through Accounting for the Case of Beer." Journal of International Economics, vol. 76, pp. 14-32.

[52] Hendel, I. and A. Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." Econometrica. Vol. 74, No. 6: pp. 1637-1673.

[53] Hill, James R., Jeff Madura and Richard A. Zuber, "The Short-run Demand for Major League Baseball." Atlantic Economic Journal, Summer, 31-35 (1982).

[54] Howard, Dennis R. and John L. Crompton, "An Empirical Review of the Stadium Novelty Effect." Sports Marketing Quarterly, Vol. 12, Issue 2 (2003).

[55] Humphreys, Brad R., "The Economic Impact of Sporting Facilities." Handbook on the Economics of Sport, Wladimir Andreff and Stefan Szymanski (Eds.), Edward Elgar Publishing (2006).

[56] Industry Profile: Beer in the United States. Multiple Issues: 2009, 2010. Datamonitor USA: New York, NY.

[57] Irani, Daraius, "Public Subsidies to Stadiums: Do the Costs Outweigh the Benefits?" Public Finance Review, 25, 238-253 (1997).

[58] Keh, Andrew. "Marlins Validate Stadium Critics' Fear." New York Times, (November, 2012).

[59] Kirby, T. and A. E. Barry. 2012. "Alcohol as a Gateway Drug: A Study of US 12th Graders." Journal of School Health: Vo. 82, Issue 8, 371-379.

[60] Knittel, C.R. and K. Metaxoglou. 2011. "In Search of the Truth: Merger Simulations using Random Coefficient Logit Models." Working Paper.

[61] Kruger, M. W. and D. Pagni. 2008. "IRI Academic Data Set Description, version 1.31." Chicago: Information Resources Incorporated.

[62] Lahman, Sean, Baseball Archive. Available at www.baseball1.com. Downloaded on January 12th, 2010.

[63] Layson, Stephen K., and M. Taylor Rhodes, "Were Major League Baseball Doubleheaders a Mistake?" UNCG Department of Economics: Working Paper 11-05, February 2011.

[64] Lemke, Robert J., Matthew Leonard and Kelebogile Tlhokwane, "Estimating Attendance at Major League Baseball Games for the 2007 Season." Journal of Sports Economics, Vol. 11, No. 3 (2010).

[65] Lentz, Bernard F. and David N. Laband, "The Impact of Intercollegiate Athletics on Employment in the Restaurant and Accommodations Industries." Journal of Sports Economics, Vol. 10 No. 4 (August 2009), pp. 351-368.

[66] Lindo, Jason M., Isaac D. Swensen and Glen R. Waddell. "Are Big-Time Sports a Threat to Student Achievement?" NBER Working Paper No. 17677 (December, 2011).

[67] Lovenheim, M.F. and D. P. Steefel. 2011. "Do Blue Laws Save Lives? The Effect of Sunday Alcohol Sales Bands on Fatal Vehicle Accidents." Journal of Policy Analysis and Management. Vol. 30, Issue 4.

[68] Marcum, John P., and Theodore M. Greenstein, "Factors Affecting Attendance in Major League Baseball: II. A Within-Season Analysis." Sociology of Sport Journal, 2, 314-322 (1985).

[69] McConnell, Bob, "The Dope Book 1942, 1948-1985: A Subject Index." Society for American Baseball Research (SABR), Research Guide #13, Feb. 1991.

[70] McCubbin, Donald. "Health Benefits of Alternative PM2.5 Standards." Report Prepared for: American Lung Association, Clear Air Task Force and Earthjustice. (July, 2011).

[71] McDonald, Mark and Daniel Rascher, "Does Bat Day Make Cents? The Effect of Promotions on the Demand for Major League Baseball." Journal of Sport Management, 14, 8-27 (2000).

[72] McGregor, Glenn R. "Basic Meteorology." Air Pollution and Health, Holgate, S. T., Koren, H. S., Samet, J. M., & Maynard, R. L. (Eds.), Academic Press (1999).

[73] Merlo, Lisa J. and Jisu Hong, Linda B. Cottler, "The Association between Alcohol-related Arrests and College Football Game Days." Drug and Alcohol Dependence 106 (2010) 69-71.

[74] Morgenstern, M., B. Isensee, J. D. Sargent and R. Hanewinkel. 2011. "Exposure to alcohol advertising and teen drinking." Preventative Medicine: Vol. 52, Issue 2, pp. 146-151.

[75] National League Green Book. The Sporting News, annual issues: 1986-2009.

[76] Nevo, A. 2000. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." Journal of Economics and Management Strategy, vol. 9, no. 4, pp. 513-58.

[77] Nevo, A. 2001. "Measuring Market Power in the Ready-To-Eat Cereal Industry." Econometrica, vol. 69, no. 2, pp. 307-42.

[78] Nevo, A. 2003. "New Products, Quality Changes and Welfare Measures Computed from Estimated Demand Systems." The Review of Economics and Statistics, Vol. 85, No. 2, pp. 266-275.

[79] Nevo, A. 2011. "Empirical Models of Consumer Behavior." Annual Review of Economics, 2:51-75.

[80] O'Keefe, J. H., K. A. Bybee and C. J. Lavie. 2007. "Alcohol and Cardiovascular Health: The Razor-Sharp Double-Edged Sword." Journal of American College of Cardiology: Vol. 50, No. 11, 1009-1014.

[81] Pedrelli, P., Bitran, S., Shyu, I., Baer, L., Guidi, J., Tucker, D. D., Vitali, M., Fava, M., Zisook, S. and Farabaugh, A. H. 2011. "Compulsive Alcohol Use and Other High-Risk Behaviors among College Students." The American Journal on Addictions: 20: pp. 14-20.

[82] Petrin, A. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." Journal of Political Economy, pp. 705-729.

[83] Popovici, I., J.F. Homer, H. Fang, and M.T. French. 2012. "Alcohol Use and Crime: Findings from a Longitudinal Sample of U.S. Adolescents and Young Adults." Alcoholism: Clinical and Experimental Research. Vol. 36, Issue 3, pp. 5532-543.

[84] Porter, Philip K. "Mega-Sports Events as Municipal Investments: A Critique of Impact Analysis." Sports Economics: Current Research, John L. Fizel, Elizabeth Gustafson and Lawrence Hadley (Eds.), Praeger Publishers (1999).

[85] Rees, Daniel I. and Kevin T. Schnepel, "College Football Games and Crime." Journal of Sports Economics, Vol. 10 No. 1 (February 2009), pp. 68-87.

[86] Rehm, J., Shield, K. D., Joharchi, N. and Shuper, P. A. 2012. "Alcohol consumption and the intention to engage in unprotected sex: systematic review and meta-analysis of experimental studies." Addiction: 107, pp. 51-59.

[87] Retrosheet Game Logs. Available at www.retrosheet.org. Downloaded on July 16th, 2010.

[88] Roerecke, Michael and Jurgen Rehm. 2012a. "The cardioprotective association of average alcohol consumption and ischaemic heart disease: a systematic review and meta-analysis." Addiction: 107; 1246-1260.

[89] Roerecke, Michael and Jurgen Rehm. 2012b. "Alcohol Intake Revisited: Risks and Benefits" Current Atherosclerosis Reports: Vol. 14, Issue 6, 556-562.

[90] Rojas, C., E.B. Peterson. 2008. "Demand for differentiated products: Price and Advertising evidence from the U.S. beer market." International Journal of Industrial Organization, 26: 288-307.

[91] Rojas, C. 2008. "Price Competition in U.S. Brewing." Journal of Industrial Economics, Vol. 56, Issue 1, pp.1-31.

[92] Ruhm, C. J., A.S. Jones, T.K. Greenfield, J.V. Terza, R.S. Pandian, and K.A. McGeary. 2011. "What U.S. Data Should be Used to Measure the Price Elasticity of Demand for Alcohol?" NBER Working Paper.

[93] Rysman, M. 2004. "Competition Between Networks: A Study of the Market for Yellow Pages." Review of Economic Studies, 71, pp. 483-512.

[94] Sacks, Jason D., Lindsay W. Stanek, Thomas J. Luben, Douglas O. Johns, Barbara J. Buckley, James S. Brown and Mary Ross. "Particulate Matter—Induced Health Effects: Who is Susceptible?" Environmental Health Perspectives, 119(4), (April, 2011), pp. 446-454.

[95] Shah, Prakesh, S. and Taiba Balkhair. "Air Pollution and Birth Outcomes: A systematic review." Environment International, 37 (2011) 498-516.

[96] Siegfried, John J., and Jeff D. Eisenberg, "Measuring and Forecasting Demand: A Case Study in Baseball." Business, January-February, 34-41 (1980a).

[97] Siegfried, John J., and Jeff D. Eisenberg, "The Demand for Minor League Baseball." Atlantic Economic Journal, 8, 56-69 (1980b).

[98] Siegfried, John and Andrew Zimbalist, "The Economics of Sports Facilities and Their Communities." Journal of Economic Perspectives, Vol. 14, No. 3 (Summer, 2000), pp. 95-114.

[99] Siegfried, John and Andrew Zimbalist, "A Note on the Local Economic Impact of Sports Expenditures." Journal of Sports Economics, Nov 2002; vol. 3: pp. 361-366.

[100] Siegfried, John and Andrew Zimbalist, "The Economic Impact of Sports Facilities, Teams and Mega-Events." Australian Economic Review (2006) vol. 39, no. 4, pp. 420-7.

[101] Snipes, D. J. and E. G. Benotsch. 2013. "High-risk cocktails and high-risk sex: Examining the relation between alcohol mixed with energy drink consumption, sexual behavior and drug use in college students." Addictive Behaviors: Vol. 38, Issue 1, pp. 1418-1423.

[102] Temple, N. J. 2012. "What Are the Health Implications of Alcohol Consumption?" Nutrition and Health: Strategies for Disease Prevention. Humana Press (Humana Press: 3rd Edition). Temple, Norman J.; Wilson, Ted; Jacobs, Jr., David R. (Eds.); pp. 323-333.

[103] Total Baseball: The Ultimate Encyclopedia of Baseball, J. Thorn and P. Palmer (Eds.), third edition (1993).

[104] Weinmayr, Gudrun, Elisa Romeo, Manuela De Sario, Stephan K. Weiland and Francesco Forastiere. "Short-term effects of PM_{10} and NO_2 on Respiratory Health among Children with Asthma or Asthma-like Symptoms: A Systematic Review and Meta-Analysis. Environmental Health Perspectives, 118, (April 2010), pp. 449-457.

[105] Wooldridge, Jeffery M., Introductory Econometrics: A Modern Approach. South-Western: 2003.

[106] World Health Organization. "WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Summary of risk assessment." (2005).

APPENDIX A

DATA APPENDIX FOR CHAPTER 1

As a check on the validity of the Retrosheet daily attendance figures from 1938-2009, we aggregate the Retrosheet daily attendance figures for each year and then compare them with Lahman's annual MLB attendance figures.[1] Figure 6 reports the percent difference between the two series from 1938 to 2008.[2] From 1950 onwards the two measures of aggregate attendance are very close, with the largest error occurring in 1952 when the Retrosheet aggregate attendance was 5% below Lahman's attendance figure. From 1938 to 1949 the Retrosheet aggregate attendance is sometimes below and sometimes above Lahman's but reasonably close except in 1949 when the Retrosheet attendance was 34% below Lahman's attendance figure. Even though the aggregate attendance for the Retrosheet data does not match Lahman's data as closely in the 1938 to 1949 period as it does in the post 1950 data, we include this earlier period because it is rich in doubleheaders.

Recall that when we use the word doubleheader we mean single-priced doubleheader because such an event offers fans an opportunity to see two games for the price of one. Separate-admission doubleheader games offer no such incentive; they are merely two games held on the same calendar day. To distinguish between these two types of doubleheaders in the Retrosheet game logs we had to make some simplifying assumptions and check their validity.

[1] For more on Sean Lahman's Baseball Archive, see www.baseball1.com.

[2] A similar plot was obtained when we compared the Retrosheet data to the total attendance data from two other sources: baseball-reference.com—made available by Rodney Fort at www.rodneyfort.com—and Total Baseball (1993).

For the doubleheaders with attendance figures there are 4 different ways in which the game logs record attendance. The first and largest group (8,837 observations), record a single attendance figure for the doubleheader. We assume these values represent the attendance of single-priced doubleheaders. The second group (658 observations), record the attendance for the first game with a positive value and the attendance for the second game with a different positive value. We assume such paired observations represent separate-priced doubleheader where fans were charged separate admissions for each game. The third group (537 observations), record the attendance for the first game with a positive value and the attendance for the second game with the same positive value. We assume that such observations redundantly entered the total attendance of a single-priced doubleheader. The fourth group (378 observations), have a positive attendance figure and a missing value for either the first or second game. We assumed that the positive attendance figure represented a single-price doubleheader. Thus we have 10,032 doubleheaders with attendance figures. We assume that 658/10,032=6.56% of these are separate-priced doubleheaders and 9,374/10,032=93.44% of these are single-priced doubleheaders.

There were also 2,115 doubleheader observations with missing attendance data for both the first and second games. These come mostly from the 1920-1937 sub-period. For our analysis using annual observations, we used team specific weights—the proportion of doubleheaders with attendance figures by team—to determine how many of the 2,115 doubleheader observations with missing attendance data represented single versus separate-priced doubleheader games. In our analysis of daily data, of course, we drop all observations for both single games and doubleheaders when there are missing observations.

We investigated the validity of our data assumptions regarding doubleheaders by contacting David Smith, President of the Retrosheet Board of Directors. He informed us that our data assumptions regarding single-priced doubleheaders were correct but mentioned the possibility of attendance inaccuracies prior to the 1980s. As an additional check, we used scheduling data from ESPN on doubleheader games from 2002 until 2009.[3] For these 216 games, we matched each doubleheader with press releases or news stories from the archives available on MLB.com. Given our data assumptions, the Retrosheet data correctly distinguished between single-priced and separate-priced doubleheaders in every case.

Lastly, we investigated a number of separate-priced doubleheader games in conjunction with an article on holiday doubleheaders by Bevis (2004). Specifically, Bevis (2004) offers a brief discussion on which teams' scheduled separate-priced holiday doubleheader games from 1934 until 1958.[4] For these teams, we were able to investigate 11 games using the Retrosheet data and given our data assumptions found that 2 games were possibly incorrectly recorded as single-priced doubleheaders, 2 games had missing attendance figures and the remainder indicated separate-priced doubleheaders. Thus in total, we believe that our data assumptions are substantially correct and offer a first attempt at separating single-priced and separate-priced doubleheaders.

In estimating equation (1), we used a collection of individual regular season game logs sorted by home team from retrosheet.org.[5] The following edits were made to the 1920 to 2009 sample. First, we dropped all attendance values which exceeded 100,000; as a result 4 observations were dropped. Next, we dropped all games played

[3] For the ESPN data on scheduled doubleheaders, see http://sports.espn.go.com/mlb/stats/doubleheaders.

[4] According to Bevis (2004), such teams were the exception rather than the rule during this period.

[5] The information used here was obtained free of charge from and is copyrighted by Retrosheet. Interested parties may contact Retrosheet at www.retrosheet.org.

in the month of March; as a result 45 observations were dropped. Lastly, we dropped all triple header games; as a result 1 observation was dropped.

After these simple corrections, the data set contained 151,323 observations. From this total, there are 127,008 observations for single games, 12,158 observations for the first game associated with a doubleheader and 12,157 observations for the second game associated with a doubleheader. Conditional on attendance being positive, there are 115,249 observations for single games, 1,328 observations for the first game associated with a doubleheader and 9,905 observations for the second game associated with a doubleheader. Conditional on attendance being zero, there are 0 observations for single games, 8,467 observations for the first game associated with a doubleheader and 0 observations for the second game associated with a doubleheader. Conditional on attendance being missing, there are 11,759 observations for single games, 2,363 observations for the first game associated with a doubleheader and 2,252 observations for the second game associated with a doubleheader.

As mentioned in the data section, the way in which attendance was recorded can be categorized into the aforementioned 4 groups. However, there are a number of exceptions. Table 28 outlines and accounts for all the ways the attendance figures associated with doubleheader games were recorded in the Retrosheet game logs. Most observations for doubleheader games appear in pairs where the first observation corresponds to all outcomes during the first game and the second observation corresponds to all outcomes during the second game.[6] In Table 28, such observations are referred to as paired observations. For example, there are 8,459 observations where attendance

[6] Approximately 64 doubleheader games were recorded in pairs but the second game associated with a doubleheader came first followed by the first game. These observations were rearranged manually.

equals 0 for the first game and equals some positive value for the second game of a doubleheader. The few remaining doubleheader games do not appear in pairs and represent either just the first game with no corresponding second game or just the second game with no corresponding first game. These games are called non-paired observations in Table 28.

For most of the data, the total attendance associated with a doubleheader game is recorded by the second game. This would include the following:

- 8,459 observations where attendance equals 0 for the first game and equals some positive value for the second game.

- 537 observations (out of 1,195) where attendance for the first game is positive and equals the attendance for the second game. We assumed that such observations redundantly entered the total attendance of a single-priced doubleheader; as such, these games remained as doubleheaders.

- 248 observations where attendance equals missing for the first game and equals some positive value for the second game.

However before dropping the first games associated with these observations we re-coded the start time–the night game dummy variable–so the second game matched the start time of the first game. This start-time correction imposed 651 real changes (20 of which set equal to missing) for the 8,459 observations where attendance equals 0 for the first game and equals some positive value for the second game. It imposed 22 real changes for the 537 observations (out of 1,195) where attendance for the first game is positive and equals the attendance for the second game. Lastly, it imposed 5 real changes (1 equal to missing) for the 248 observations where attendance equals missing for the first game and equals some positive value for the second game. After these changes, the observations corresponding to the first games of a doubleheader were dropped. Thus, 8,459 observations were deleted from the first sub-sample, 537 from the second and 248 from the third.

For a portion of the remaining observations, we have the following sub-samples:

- 130 observations where attendance for the first game is positive and the attendance for the second game is missing. Such observations were re-coded as a second game of a doubleheader.

- 6 observations for the split-stadium doubleheader games between the New York Yankees and New York Mets in 2000, 2003 and 2008. Press releases from MLB.com suggested that these were all separate-priced doubleheader games. Thus, these observations were re-coded as single games.

For the last remaining sub-samples, we have the following:

- 658 observations (out of 1,195) where attendance for the first game is positive, the attendance for the second game is positive, and both values were not equal. We assumed such observations represented a doubleheader game where fans were charged separate admissions for each game; as such, these games were re-coded as single games. Thus, 658 real changes were made.

- 1 observation where attendance for the first game is zero and there are no recorded second games associated with these particular doubleheaders. These observations were dropped; thus, 1 real change was made.

- 7 observations where attendance for the first game was zero and the attendance for the second game was missing. These observations were dropped; thus, 7 real changes were made.

- 2,115 observations where attendance for both the first and second game was missing. These observations were dropped; thus, 2,115 real changes were made.

In this newly constructed data set, there are a total of 137,704 observations. Of this total, there are 128,330 observations for single games and 9,374 observations for doubleheader games. The gain of 1,322 observations for single games is accounted for by the changes made to the 658 observations (out of 1,195) where attendance for the first game is positive, the attendance for the second game is positive, and both values were not equal. Recall, these observations were re-coded as single games and thus would represent a net-gain of 2 658 observations. In addition, the 6 split-stadium games between the New York teams were re-coded as single games.

To account for the loss of 2,783 doubleheaders, recall our initial sample contained 151,323 observations, 12,157 of which were for the second game associated with a doubleheader. Of this 12,157 total, 9,905 corresponded to the second game of a doubleheader conditional on attendance being positive. Also, we had 0 observations for the second game associated with a doubleheader conditional on attendance being zero and 2,252 observations where attendance was missing. Further, after all the changes, the attendance levels associated with the second game of a doubleheader are always positive. This creates the loss of 2,783 observations, 2,252 of which are explained by the loss associated with no longer having missing attendance figures for the second game of a doubleheader. The remaining 531 can be explained by the loss of 658 observations, the gain of 130 observations and the loss of 3 observations.

The loss of 658 comes from the observations where attendance for the first game was positive, the attendance for the second game was positive, and both values were not equal. Recall these were re-coded as single games. The gain of 130 comes from the 130 observations where attendance for the first game was positive and the attendance for the second game was missing. Recall these observations were re-coded as second games of a doubleheader. Lastly, the loss of 3 comes from the second games of the split-stadium doubleheader games between the New York Yankees and New York Mets. Recall, these observations were re-coded as regular games.

For the regression analysis, we restricted the data from 1938 to 2009 due to the high frequency of missing observations on attendance from 1920 to 1937. Specifically from 1920 until 1937, the smallest proportion of missing values on attendance was about 35% (1924) and the largest was 77% (1934). Thus, dropping all observations prior to 1938 resulted in a loss of 17,381 observations. As a result, the total number of observations declined from 137,704 to 120,323.

For the sample sizes in our regression results, the reported total of 118,477 observations in the daily regression results reported in Table 2 is explained by taking the total number of 120,323 observations and then subtracting the 1,149 missing observations on attendance, subtracting the 728 missing observations on the dummy variable for night games and adding the 31 observations where night games and attendance are both missing in order to avoid double-counting.

This portion offers a brief discussion on the standard controls which were included in all daily regressions discussed in sections 1.5 and 1.6 of the paper. Specifically, the results in Table 29 correspond to the set of controls included in the regression results presented in Table 1 of the text. Almost all of the controls are highly statistically significant.

The results in Table 29 indicate that fans enjoy higher quality home and away teams. In particular, home team quality both in terms of current and last season winning percentage increases home team daily attendance by about 400 and 200, respectively. In addition, home team daily attendance increases if the visiting team is higher quality both in terms of current season and last season performance. Opening day illustrates the initial interest in all teams at the start of the season and has a substantially large impact with an estimated increase in home team daily attendance of nearly 25,000. Also, the addition of a new stadium increases home team attendance by about 7,500 although this honeymoon effect is limited to just one year—see Howard et al. (2003) and Clapp et al. (2005) for some recent findings on the honeymoon effects of new stadiums.

Home team daily attendance tends to be higher when games are played during fan leisure hours. Specifically, night games increase daily home team attendance by about 2,500. For holiday games, Labor Day increases daily home team attendance the most (8,400) followed by Memorial Day (8,000) and lastly Independence Day (5,000).[7] Weekend days like Friday, Saturday and Sunday have higher attendance levels than weekdays. Lastly, vacation and summer time months of June, July and August tend to also have higher attendance.

Lastly, Tables 30 and 31 present regression results for two alternative specifications discussed in section 1.6 of the paper.

[7] Data on the specific calendar dates for each of the holidays was taken from www.timeanddate.com.

Table 28. How the Retrosheet data record Attendance and Doubleheader Games (Data 1920-2009).

(1) *First Game Attendance*	(2) *Second Game Attendance*	(3) *Number of Observations*
I. Paired Observations		
1st Game = 0	2nd Game > 0	8,459
1st Game > 0	2nd Game > 0	1,195
1st Game > 0	2nd Game = 0	0
1st Game = 0	2nd Game = 0	0
1st Game = missing	2nd Game > 0	248
1st Game > 0	2nd Game = missing	130
1st Game = missing	2nd Game = missing	2,115
1st Game = 0	2nd Game = missing	7
1st Game = missing	2nd Game = 0	0
II. Non-paired Observations		
	Solo 2nd Game > 0	3
	Solo 2nd Game = 0	0
	Solo 2nd Game = missing	0
Solo 1st Game > 0		3
Solo 1st Game = 0		1
Solo 1st Game = missing		0

Table 29. Full Sample Results for Standard Controls.

	(1)	(2)
	One Lag and Lead	*Additional Lags and Leads*
Home Team Season Win %	416***	416***
	(30)	(30)
Lag Home Team Season Win %	221***	221***
	(30)	(30)
Away Team Season Win %	195***	195***
	(12)	(12)
Lag Away Team Season Win %	141***	140***
	(8)	(8)
Night Game	2,580***	2,591***
	(433)	(433)
Opening Day	19,518***	19,337***
	(1,253)	(1,250)
Memorial Day	8,042***	8,006***
	(864)	(861)
Independence Day	4,848***	4,805***
	(951)	(948)
Labor Day	8,335***	8,372***
	(810)	(802)
New Stadium	7,468***	7,480***
	(816)	(814)
Tuesday	-99	-166
	(237)	(235)
Wednesday	77	-117
	(215)	(212)
Thursday	109	83
	(216)	(207)
Friday	3,980***	4,033***
	(377)	(375)
Saturday	7,888***	7,648***
	(433)	(428)
Sunday	8,507***	8,366***
	(603)	(595)
May	1,931***	2,032***
	(233)	(234)
June	4,987***	5,137***
	(420)	(418)
July	6,150***	6,299***
	(483)	(478)
August	5,394***	5,544***
	(488)	(487)
September	662*	712**
	(332)	(333)
October	930*	705

	(549)	(548)

Notes: Robust standard errors in parentheses, clustered by home team; * significant at 10%; ** significant at 5%; *** significant at 1%. Above are the standard controls included in the regression results shown in Table 1. Home team and year dummies were also included. The dependent variable is home team daily attendance.

Table 30. Daily Results after Adding Home Team and Year Interactions for Full-sample and 3 Sub-samples.

	(1)	(2)	(3)	(4)
	1938-2009	*1938-1947*	*1953-1984*	*1987-2009*
Doubleheader	3,760***	4,354***	4,246***	-2,065***
	(454)	(661)	(575)	(702)
S 1 Day Prior S	-639***	-252	-376**	-862***
	(118)	(235)	(160)	(121)
S 2 Day Prior S	-147*	401*	-62	-257***
	(76)	(224)	(136)	(70)
S 3 Day Prior S	-171**	63	-325**	-251**
	(81)	(194)	(147)	(121)
S 4 Day Prior S	10	-414	134	-184*
	(96)	(274)	(126)	(107)
S 5 Day Prior S	-463***	-250	-375***	-241**
	(73)	(169)	(102)	(96)
S 6 Day Prior S	-180**	-36	-22	-227**
	(78)	(174)	(111)	(103)
S 7 Day Prior S	-229**	145	-306**	-419***
	(95)	(278)	(117)	(123)
S 1 Day After S	-835***	-909***	-818***	-646***
	(99)	(217)	(107)	(95)
S 2 Day After S	-197***	94	-188	-203***
	(71)	(194)	(116)	(61)
S 3 Day After S	-432***	-16	-323**	-672***
	(87)	(192)	(133)	(88)
S 4 Day After S	151**	-3	151	162**
	(67)	(174)	(104)	(78)
S 5 Day After S	168***	172	234**	-40
	(55)	(164)	(93)	(96)
S 6 Day After S	-114	-104	-109	-413***
	(91)	(179)	(95)	(81)
S 7 Day After S	-557***	22	-417***	-391***
	(59)	(173)	(76)	(111)
S 1 Day Prior D	-3,910***	-1,318**	-2,263***	-1,691***
	(331)	(460)	(363)	(397)
S 2 Day Prior D	-260	296	230	-1,576***
	(283)	(280)	(352)	(363)
S 3 Day Prior D	-344**	433	-362	-543
	(148)	(331)	(220)	(437)
S 1 Day After D	-1,296***	-1,164**	-1,345***	-1,429***
	(224)	(439)	(206)	(407)

S 2 Day After D	-25	-85	-220	-748**
	(183)	(349)	(219)	(348)
S 3 Day After D	-233	178	-239	-411
	(191)	(220)	(192)	(426)
D 1 Day Prior D	-4,657***	-2,643***	-3,116***	683
	(391)	(622)	(977)	(2,637)
D 2 Day Prior D	-800**	350	-1,242**	1,609
	(385)	(614)	(502)	(3,019)
D 1 Day After D	-408	-389	-1,835**	4,450
	(547)	(1,037)	(727)	(2,967)
D 2 Day After D	-666*	-215	-1,318**	-274
	(377)	(613)	(490)	(2,626)
D 1 Day Prior S	-881**	857	-466	198
	(331)	(605)	(435)	(702)
D 1 Day After S	1,192***	1,364**	476	1,932**
	(390)	(581)	(449)	(751)
Observations	118477	9183	49007	51607
R-squared	0.7223	0.6294	0.6178	0.7267

Notes: Robust standard errors in parentheses, clustered by home team; * significant at 10%; ** significant at 5%; *** significant at 1%. Standard controls discussed in the text are included and the interaction of home team and year are also included. Section 4 and footnotes 11 and 12 offer the proper interpretation of all point-estimates shown above. The dependent variable is home team daily attendance.

Table 31. Daily Results after Dropping Sellouts for Full-sample and 3 Sub-samples.

	(1)	(2)	(3)	(4)
	1938-2009	*1938-1947*	*1953-1984*	*1987-2009*
Doubleheader	3,173***	3,435***	4,092***	-2,123**
	(422)	(725)	(565)	(893)
S 1 Day Prior S	-477***	-270	-227	-768***
	(117)	(208)	(173)	(117)
S 2 Day Prior S	-104	413**	24	-228***
	(71)	(184)	(142)	(62)
S 3 Day Prior S	-117	60	-247	-276**
	(88)	(189)	(149)	(130)
S 4 Day Prior S	88	-236	289**	-154
	(94)	(246)	(130)	(116)
S 5 Day Prior S	-462***	-140	-323***	-295***
	(74)	(176)	(114)	(93)
S 6 Day Prior S	-148*	8	125	-328***
	(79)	(197)	(112)	(99)
S 7 Day Prior S	-193*	61	-167	-393***
	(96)	(251)	(116)	(123)
S 1 Day After S	-768***	-951***	-669***	-629***
	(113)	(219)	(112)	(104)
S 2 Day After S	-210**	-38	-99	-221***
	(83)	(200)	(132)	(69)
S 3 Day After S	-364***	-33	-252*	-673***
	(94)	(177)	(128)	(94)
S 4 Day After S	222***	-31	253**	202**
	(71)	(161)	(113)	(90)
S 5 Day After S	156**	212	325***	-92
	(73)	(158)	(110)	(98)
S 6 Day After S	-102	-95	51	-528***
	(95)	(151)	(99)	(92)
S 7 Day After S	-484***	135	-270**	-380***
	(74)	(179)	(102)	(116)
S 1 Day Prior D	-3,987***	-1,538***	-2,479***	-2,770***
	(317)	(386)	(381)	(464)
S 2 Day Prior D	-468	257	-39	-2,272***
	(352)	(307)	(398)	(453)
S 3 Day Prior D	-617**	8	-753***	-1,014*
	(248)	(237)	(260)	(507)
S 1 Day After D	-1,625***	-1,380***	-1,544***	-2,396***
	(274)	(382)	(198)	(523)
S 2 Day After D	-343	-178	-486**	-1,791***

	(256)	(352)	(217)	(488)
S 3 Day After D	-575**	-31	-592**	-1,518**
	(261)	(232)	(231)	(582)
D 1 Day Prior D	-4,122***	-1,585**	-3,342***	201
	(475)	(650)	(1,022)	(3,023)
D 2 Day Prior D	-1,032**	-14	-1,483***	212
	(443)	(596)	(513)	(2,919)
D 1 Day After D	-223	0.479	-2,206***	4,336
	(597)	(1,067)	(691)	(3,386)
D 2 Day After D	-630	-263	-1,283**	-1,440
	(408)	(547)	(611)	(2,121)
D 1 Day Prior S	-787**	1,092	-338	-558
	(387)	(628)	(438)	(869)
D 1 Day After S	1,165***	896*	597	1,599**
	(391)	(485)	(451)	(725)
Observations	115313	8964	48456	49400
R-squared	0.579	0.582	0.507	0.554

Notes: Robust standard errors in parentheses, clustered by home team; * significant at 10%; ** significant at 5%; *** significant at 1%. Standard controls discussed in the text are included. Section 4 and footnotes 11 and 12 offer the proper interpretation of all point-estimates shown above. The dependent variable is home team daily attendance.

Capacity Data: Capacity data came from a variety of sources including: baseball-almanac.com (1920-1948), the Dope Book (1949-1985), the AL Red Book and NL Green Book (1986-2009) and ballparks.com. Specifically, ballparks.com was used for the following teams and years: Cleveland Indians (1920-1948), Philadelphia Phillies (1920-1937), St. Louis Cardinals (1920-1948), St. Louis Browns (1920-1948), Arizona Diamondbacks (2009), Toronto Blue Jays (2009), Washington Nationals (2009). Also, for a complete description of all statistics contained in the Dope Book series see McConnell (1991).

Figure 6. Percent Difference in Attendance Data Reported by Retrosheet and Lahman (Data 1938-2008).